

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**DEVELOPMENT OF AN
ARTIFICIAL-INTELLIGENCE-DRIVEN
PRODUCT DESIGN EVALUATION
MODEL USING MULTI-MODAL
DATA**

JING LUO

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

School of Design

**Development of an Artificial-Intelligence-Driven Product Design
Evaluation Model Using Multi-modal Data**

Jing LUO

A thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

August 2024

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

_____ (Signed)

_____ **Jing LUO** _____ (Name of student)

Abstract

This thesis systematically explores the potential and practical applications of artificial intelligence (AI) in enhancing product design evaluation. As AI technology advances, its integration with design processes offers new approaches to improve efficiency, innovation, and decision-making in product design. This study focuses on developing and validating a multimodal AI-assisted product design evaluation system, leveraging deep learning algorithms that can process and analyze complex multimodal data, images and text, to provide comprehensive evaluation metrics.

The contribution of this study is the establishment of a framework that integrates AI with traditional product design evaluation practices. The research is organized around five interrelated studies, each of which addresses a specific application of AI in product design—from data collection and model development to practical validation and human-machine collaboration. The results show that AI can significantly improve the objectivity and efficiency of design evaluation, especially when dealing with large-scale and multidimensional datasets.

AI's ability to synthesize large amounts of design-related information has been shown to significantly enhance the decision-making process, allowing for faster iterations and more informed adjustments during the design phase.

However, the study also acknowledges the challenges of AI in its application, especially in dealing with highly subjective design aspects such as aesthetics and user experience, where human insight remains indispensable.

In addition, this paper proposes a human-machine collaborative model as an ideal approach for design evaluation, combining the analytical advantages of artificial intelligence with human creativity and critical judgment. This model not only improves the reliability of design evaluation, but also promotes innovation by integrating different perspectives.

In terms of practice, the developed AI-assisted evaluation system has been tested and validated through empirical methods, showing usability and effectiveness results in real-world design scenarios. This study contributes to the theoretical and practical understanding of the application of artificial intelligence in design and suggests avenues for future innovation and ethical considerations for the integration of artificial intelligence technology in the design industry.

Overall, this study not only highlights the transformative potential of artificial intelligence in product design evaluation, but also lays the foundation for future advances to promote smarter, more efficient, and user-centered design practices.

Publications

Luo, J., Wu, L., & Luximon, Y. (2024). A systematic literature review of sensory-related sustainable product design. *Environmental Development*, 100993.

Luo, J., Feng, L., & Luximon, Y. (2024). Impacts and opportunities of deep learning in product design process: a comprehensive survey. *Journal of Engineering Design*, 1-28.

Luo, J., Luximon, Y., Zhan, W., & Chen, X. (2020). An Ergonomic Solution for Hand Rehabilitation Product Design for Stroke Patients. *Proceedings of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I 22* (pp. 325-334). Springer International Publishing.

Luo, J., Luximon, Y., & Ni, X. (2021). A Conceptual Design of a Modular Multifunctional Nursing Bed for Moderately Disabled Patients. *Proceedings of the AHFE 2021 Virtual Conferences on Creativity, Innovation and Entrepreneurship, and Human Factors in Communication of Design, July 25-29, 2021, USA* (pp. 128-135). Springer International Publishing.

Zhong, Y., Qu, J., Zheng, K., Luximon, Y., & Luo, J. (2024, June). Improving product design efficiency through integrated AI tools: an empirical study. *Proceedings of the 15th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences*. AHFE Open Access.

Acknowledgments

Throughout the research and writing process of this doctoral dissertation, I have been fortunate to receive help and support from many individuals. I extend my heartfelt thanks to each one of them.

First and foremost, I would like to express my special gratitude to Professor Yan Tina Luximon, who has provided me with selfless guidance and patient instruction throughout the entire research process. Professor Yan Tina Luximon has not only been a tremendous academic mentor but also offered me support and care in my personal life, keeping me motivated and confident when facing difficulties and challenges.

I am also grateful to all the team members of Asian Ergonomics Design Lab for the inspiring and enjoyable collaboration. Thank you for the valuable suggestions and technical support in experimental design, data collection, and thesis writing.

I must also thank my family, for their unconditional love and support. It is your encouragement that has allowed me to persevere and complete this academic journey.

Lastly, thanks to all colleagues and friends who participated in and supported my research work, especially Mr. Pengzai and Ms. Xiaoqiao. Without your assistance, this thesis could not have reached its current stature.

My deepest thanks go out to everyone who has provided me with help and encouragement!

Contents

List of Figures	1
List of Tables	5
Chapter 1. Introduction	8
1.1 Research Background and Significance	8
1.2 Theoretical Foundation	19
1.2.1 Artificial Intelligence Theory	19
1.2.2 Product Design Theory	20
1.2.3 Human-computer Collaboration Theory	21
1.3 Research Objectives	23
1.4 Overview of the Thesis Structure	25
Chapter 2. Literature Review	29
2.1 Current Status of AI Application in Product Design Evaluation	29
2.1.1 Current Status and Impact of AI in Product Design	29
2.1.2 Challenges and Future Prospects of AI Applications	32
2.2 Limitations of Existing Research	35
2.2.1 Limitations of Design Evaluation Studies	35
2.2.2 Challenges of Data Quality, AI Explainability, and Evaluation Framework	38
2.3 Potential of Multimodal Data and Deep Learning in Product Design Evaluation	40
2.3.1 Improving the Comprehensiveness and Accuracy of Design Evaluation	40
2.3.2 Promote Product Innovation and Meeting User Needs	41
2.3.3 Implementation of Real-time Evaluation and Iterative Optimization ...	41
2.3.4 Enhance Interdisciplinary Collaboration and Decision Support	42
2.4 Research Gaps and Research Questions	43
Chapter 3. Research Methodology	48
3.1 Research Methods	48
3.1.1 Qualitative Methods	50
3.1.1.1 Focus Group	50
3.1.1.2 Grounded Theory	52
3.1.2 Quantitative Methods	53
3.1.2.1 Questionnaires	53
3.1.2.2 Natural Language Processing (NLP)	56
3.1.2.3 Computer Vision (CV)	59
3.1.2.4 Multimodal Neural Network	61
3.1.2.5 Experiments	63
3.1.3 Mixed Methods: Questionnaire& Structured Interview	65
3.1.4 Reasons for the Mixed Methods	69
3.2 The Methodology Framework	71
3.2.1 Definition	72

3.2.2 Data	74
3.2.3 Modeling	76
3.2.4 Validation	77
3.3 Summary of Chapter Three	81
Chapter 4. Study 1- Design Tool Requirements in the Age of AI	85
4.1 Introduction	85
4.2 Methods and Results	87
4.2.1 Focus Group Discussion	87
4.2.2 Grounded Theory Analysis	92
4.2.2.1 Open Coding	93
4.2.2.2 Axial Coding	98
4.2.2.3 Selective Coding	102
4.2.3 Summary of Results	103
4.3 Summary and Discussions	105
Chapter 5. Study2- AI-assisted Design Evaluation Data Requirements	111
5.1 Introduction	111
5.2.1 Questionnaire Design	114
5.2.2 Data Collection	114
5.2.3 Data Analysis	115
5.3 Results	116
5.3.1 Sample Characteristics	116
5.3.2 Designers' Understanding and Current Usage of AI Tools	117
5.3.3 AI Design Evaluation Needs Analysis and Challenges	125
5.3.4 Expectations for AI Evaluation Tools	130
5.4 Summary and Discussions	140
Chapter 6. Study3 Part I- Construction of Multimodal Product Design Datasets	144
6.1 Introduction	144
6.2 Methods	146
6.2.1 Datasets Collection	146
6.2.2 Datasets Processing Methods	147
6.2.3 NLP Text Analysis	149
6.3 Results	149
6.3.1 Database Statistical Analysis	149
6.3.2 Data Analysis by Natural Language Processing	154
6.3.3 Functional Contribution Analysis of Multimodal Data	158
6.4 Summary and Discussions	159
Chapter 7. Study3 Part II- Implementation of AI Evaluation Model	163
7.1 Introduction	163
7.2 Methods	165
7.2.1 AI Evaluation Algorithm	167
7.2.1.1 Visual Feature Extraction Module	167
7.2.1.2 Vision-based Product Rating Algorithm	170
7.2.1.3 Bert-based Product Rating Algorithm	171
7.2.1.4 Vision-Language Fusion for Product Rating Algorithm	173

7.2.2 Training and Loss Function	174
7.2.3 Performance Metric	175
7.2.4 Dataset Quality Control and Deviation Analysis	176
7.3 Results	178
7.3.1 Model Training Performance	178
7.3.2 Result Analysis	182
7.4 Summary and Discussions	184
Chapter 8. Study4- Human-machine Comparative Evaluation Experiment	189
8.1 Introduction	189
8.2 Methods	193
8.2.1 GUI Design and Website Function Implementation	193
8.2.1.1 GUI Design	194
8.2.1.2 Website Function Implementation	202
8.2.2 Human-machine Comparison Experiment Design	204
8.2.2.1 Participant Selection and Recruitment	205
8.2.2.2 Experimental Process	205
8.2.2.3 Human-machine Comparison Rating	207
8.2.3 Evaluation Indicators	208
8.3 Results	209
8.3.1 Website Usage Data Analysis	209
8.3.1.1 Participants and Sample	209
8.3.1.2 Analysis of Exports Rating Behavior	211
8.3.1.3 User Interface and Interaction Design	212
8.3.1.4 Human-machine Collaboration Evaluation	212
8.3.2 Comparative Analysis of Human-machine Ratings	215
8.4 Artificial Intelligence Algorithm Optimization	219
8.5 Summary and Discussions	225
Chapter 9. Study 5- Prototype Validation Research	231
9.1 Introduction	231
9.2 Methods	234
9.2.1 Questionnaire	234
9.2.2 Structured Interview	235
9.3 Data Analysis and Results	237
9.3.1 Questionnaire Analysis Results	237
9.3.1.1 The Limitations of AI Design Evaluation	237
9.3.1.2 Trust in AI Evaluation Models and Willingness to Use	241
9.3.1.3 Relationship between Familiarity with AI and Confidence/ Willingness in Using AI	245
9.3.1.4 Relationship between Length of Time Engaged in Design and Confidence in Using AI	248
9.3.1.5 Open Questions	250
9.3.2 Interview Coding Results	255
9.4 Summary and Discussions	258
Chapter 10. Discussion	263

10.1 General Research Results	263
10.1.1 Advantages of AI Models	266
10.1.2 Collaboration between AI and Human Designers	268
10.2 Discussion of the Challenges of AI in Subjective Evaluation Tasks and the Extension of Future Research	271
10.2.1 Application Effect of AI in Product Design Evaluation	271
10.2.2 Effectiveness of Multimodal Data Fusion	272
10.2.3 The Role of Deep Learning Technology in Improving Design Evaluation	272
10.2.4 Challenges and Future Research Directions	273
10.2.4.1 Transparency and Interpretability of Model	273
10.2.4.2 Quality and Diversity of Data	275
10.2.4.3 Discussion of Future Research Directions	275
10.3 Discussion of Exploring and Optimizing the Human-computer Collaboration Model	278
10.3.1 Improving the Human-computer Interface	278
10.3.2 Enhancing the Credibility and Interpretability of AI-assisted Evaluation	279
10.4 Limitations and Future Research Directions	281
10.4.1 Limitations	281
10.4.2 Future Research Directions	283
Chapter 11. Conclusion	286
11.1 Main Research Findings	286
11.1.1 Development and Application of Multimodal AI Evaluation Models	287
11.1.2 Potential for Improving Design Efficiency and Innovation of AI	288
11.1.3 Evaluation Model of Human-computer Collaboration	288
11.2 Theoretical and Practical Contributions	290
11.2.1 Theoretical Contributions	291
11.2.1.1 Application of Multimodal Deep Learning in Design Evaluation ...	291
11.2.1.2 Establishment of Human-computer Collaboration Evaluation Framework	292
11.2.1.3 Research on AI Explainability in Design Evaluation	292
11.2.2 Practical Significance	293
11.2.2.1 Providing Practical AI-assisted Evaluation Tools	294
11.2.2.2 Optimizing the Product Design Innovation Process	295
11.2.2.3 Promote the Application of AI Technology in the Design Industry	296
11.2.2.4 Provide Reference for Design Education	296
11.3 Summary of Conclusion	297
References	300
Appendix A. Ethical Approval	319
Appendix B. Consent form in Study 1, 2, 4 & 5	320

List of Figures

Figure 1.1 The Thesis Structure	25
Figure 3.1 The Framework of Methodology	72
Figure 4.1 Focus Group- Senior Designers	90
Figure 4.2 Focus Group- Product Design Students	92
Figure 4.3 The Coding Flowchart.....	93
Figure 4.4 The Logical Relationship of Axial Coding	99
Figure 4.5 The Relationship between Codes	103
Figure 5.1 Usage of Various AI Tools	119
Figure 5.2 The Ideal Interaction between Designers and AI	120
Figure 5.3 AI Technology Evaluation Design Project Usage	121
Figure 5.4 The Stage of Conducting Design Evaluation	125
Figure 5.5 Factors of Concern and Data Sources	127
Figure 5.6 Evaluation Data Source Word Cloud	128
Figure 5.7 Difficulties in the Evaluation Process Word Cloud	129
Figure 5.8 Reasons for not Using AI Evaluation Tools	130
Figure 5.9 Text Analysis of Other Options	131
Figure 5.10 The Interest of Designers in AI Evaluation Products	133
Figure 5.11 The Willingness of Designers to Learn about AI Evaluation Products	133
Figure 5.12 Potential Stage of AI Evaluation	134
Figure 5.13 Cross Summary of Application Stages and Advantages	135

Figure 5.14 Data Types can Assist in Conceptual Design Evaluation	136
Figure 5.15 Feedback form of Evaluation Results	138
Figure 6.1 An Example of the Basic Data	147
Figure 6.2 Data Classification	148
Figure 6.3 Data Processing	149
Figure 6.4 An Example of the Database	151
Figure 6.5 The Distribution of Scores	152
Figure 6.6 Example of Experts Rating Results	152
Figure 6.7 SRCC	153
Figure 6.8 PLCC	154
Figure 6.9 The Word Cloud of the Low Score	155
Figure 6.10 The Word Cloud of the Median Score	156
Figure 6.11 The Word Cloud of the High Score	156
Figure 6.12 The Relationship between Text Length and Score	157
Figure 7.1 The Architecture of the Multimodal Deep Learning Model	166
Figure 7.2 Vision-based Product Rating Algorithm	171
Figure 7.3 Bert-based Product Rating Algorithm	172
Figure 7.4 Vision-Language Fusion for Product Rating Algorithm	174
Figure 7.5 An Example of Training Results	183
Figure 8.1 Registration Page	194
Figure 8.2 User Permission Setting Function	195
Figure 8.3 User Interface Design	195

Figure 8.4 The User Upload Design Project Interface	196
Figure 8.5 Visualization of Rating Results	197
Figure 8.6 Expert rating process	197
Figure 8.7 Gallery	198
Figure 8.8 Machine Rating icon	199
Figure 8.9 Expert Rating icon	199
Figure 8.10 Delete Product icon	200
Figure 8.11 Backend Login	201
Figure 8.12 Backend Management	201
Figure 8.13 Website Framework	202
Figure 8.14 Experimental Process	206
Figure 8.15 Experimental Procedures	207
Figure 8.16 Number of Design Projects	210
Figure 8.17 Number of Ratings	210
Figure 8.18 Rating Sample	211
Figure 8.19 A Sample of Similar Human-machine Rating Results	213
Figure 8.20 A Sample of Average Performance	214
Figure 8.21 A Sample of Significant Differences	215
Figure 8.22 A Sample of Functional Product	217
Figure 8.23 A Sample of Non-functional Product	217
Figure 8.24 Machine Learning Training	220
Figure 8.25 Log Output	221

Figure 8.26 Training and Validation	222
Figure 9.1 Main Limitations of AI in Concept Evaluation.....	238
Figure 9.2 The Role of AI in Conceptual Design Evaluation.....	251
Figure 9.3 AI Concept Design Evaluation Expected Functions.....	251
Figure 9.4 Needs of AI Concept Design Evaluation.....	252
Figure 9.5 Suggestions for Improving AI Design Evaluation.....	253
Figure 9.6 Expectations and Suggestions for AI in Design Evaluation.....	254
Figure 11.1 The Human-computer Collaborative Evaluation Framework	289

List of Tables

Table 4.1 Focus Group Outline	88
Table 4.2 Participants- Senior Designers	89
Table 4.3 Open Codes	94
Table 4.4 Axial Codes	99
Table 4.5 Selective Codes	102
Table 5.1 Characteristics of the Participants	116
Table 5.2 Relationship between Region and AI Familiarity	118
Table 5.3 Relationship between Related Background and Familiarity with AI..	118
Table 5.4 Usage of AI Tools	119
Table 5.5 Relationship between Familiarity with AI Design Evaluation Tools and Related Indicators	122
Table 5.6 Rating of AI Accuracy, Potential Attitude, and Other Aspects by Chinese and American Designers	123
Table 5.7 Cross Analysis of Factors of Concern and Data Sources in Conceptual Design Evaluation	126
Table 5.8 Data Sources for Design Evaluation	127
Table 5.9 Difficulties in the Design Evaluation Process	128
Table 5.10 Interest and Learning Willingness of AI Evaluation	132
Table 5.11 Summary of AI Application Stages and Advantage Crossovers	134
Table 5.12 Designers' Evaluation of AI Technology in Design Evaluation	135
Table 5.13 Summary of Response Rate and Popularity Rate of Data Types	136
Table 5.14 Co-occurrence Matrix of Options of Data Types	137
Table 5.15 Summary of Response Rate and Popularity Rate of Feedback Forms	

.....	138
Table 5.16 Co-occurrence Matrix of Options for Combination Preferences	139
Table 6.1 Projects Information	150
Table 6.2 Evaluations Information	150
Table 7.1 Vgg16	168
Table 7.2 ResNet18	168
Table 7.3 Desene121	169
Table 7.4 CustomNet	169
Table 7.5 ViT	170
Table 7.6 Model Training Results of Design	179
Table 7.7 Model Training Results of Technology	179
Table 7.8 Model Training Results of Market	180
Table 7.9 Model Training Results of Investment	181
Table 7.10 Model Training Results of Media	181
Table 8.1 Results of Human-machine Rating	216
Table 8.2 Results of Human-machine Rating for Functional Products	218
Table 8.3 Results of Human-machine Rating for Non-functional Products	218
Table 8.4 Optimized Human-machine Rating Results	223
Table 8.5 Optimized Results for Functional Products	224
Table 8.6 Optimized Results for Non-functional Products	224
Table 9.1 Summary of Response Rate and Penetration Rate	239
Table 9.2 Option Co-occurrence Matrix	240

Table 9.3 AI Evaluation Trust Scale Cronbach Reliability Analysis	241
Table 9.4 AI Evaluation Trustworthiness Rating	242
Table 9.5 Cronbach Reliability Analysis Using the Willingness scale	243
Table 9.6 Willingness to Use AI Evaluation Tools	243
Table 9.7 Pearson Correlation of Each Dimension	244
Table 9.8 The Relationship between Familiarity and Using Confidence	246
Table 9.9 The Relationship between Familiarity and Using Willingness	246
Table 9.10 The Relationship between Working Experience and Using Confidence	248
Table 9.11 The Relationship between Working Experience and Using Willingness	249
Table 9.12 Open Codes of Interview	256
Table 9.13 Axial Codes of Interview	257
Table 9.14 Selective Codes of Interview	257

Chapter 1. Introduction

This chapter introduces the research background, the significance of AI-assisted product design evaluation, and the framework of this thesis. With the rapid development of AI technology, its application in product design is becoming increasingly extensive. However, there is exciting potential in design evaluation. This thesis will systematically explore the possibility of AI-assisted product design evaluation and provide a new perspective for improving design efficiency and quality. This chapter examines the theoretical foundation of this study and briefly introduces the structural design of the paper as a preparatory framework for further in-depth analysis.

1.1 Research Background and Significance

Product design evaluation is crucial to the modern product development and innovation process. The characteristic of the decision-making problem is its multi criteria nature, including factors related to the task and decision-making process (Liu&Kim, 2023). The main significance of design evaluation is to identify and select satisfactory design concepts in the early stages of product development, laying the foundation for subsequent detailed design (Xu et al., 2019). This approach enables designers and enterprises to significantly reduce time and financial expenditures while simultaneously enhancing the final product's alignment with user requirements and market demands. Studies have shown that decisions made in the early design stage often determine about 70% of the total cost of product development (Liu & Kim, 2023) . Therefore, effective design evaluation can significantly reduce the risk of product development and improve resource utilization efficiency. Designs are

typically evaluated from multiple perspectives: the functional dimension (measured by customer satisfaction), the physical dimension (assessing quality, reliability and durability), and the economic dimension (analyzing costs and investment evaluations) (Xu et al., 2019). This multi-dimensional evaluation method can comprehensively consider all aspects of the product and ensure the feasibility and competitiveness of the final design solution.

However, the design evaluation process also faces many challenges. First, the identification of evaluation criteria and the allocation of weights are two key issues. These criteria may vary from company to company and need to be adapted to the needs of customer satisfaction (Liu & Kim, 2023). In practice, expert judgment is often inconsistent in the process of identifying evaluation criteria, and it is easy to ignore existing design knowledge and experience in the early design stage (Liu & Kim, 2023). This inconsistency and lack of knowledge may lead to biased evaluation results and affect the final design decision. Another important challenge is the uncertainty and imprecision of information. In the early design stage, design information is often insufficient and imprecise. Decision makers' judgments are usually imprecise, and the confidence level in their judgments also leads to varying degrees of uncertainty (Liu & Kim, 2023). This information uncertainty increases the complexity of design evaluation and may lead to doubts about the reliability of evaluation results. In addition, design evaluation is usually a group decision-making problem involving multiple decision makers (Liu & Kim, 2023). How to effectively aggregate the judgments of different decision makers and coordinate possible differences is an important challenge in the design evaluation process. This requires the establishment of a scientific and fair decision-making mechanism to ensure the objectivity and representativeness of the evaluation results.

In the current technological environment, the application of artificial intelligence (AI) technology provides new possibilities for solving some problems in design evaluation. AI technology can help process large amounts of data and identify potential patterns and relationships, thereby providing a more objective and comprehensive analysis basis for design evaluation (Liu & Kim, 2023; Rodgers & Huxor, 1998) . For example, through machine learning algorithms, valuable experience and knowledge can be extracted from historical design cases to assist in the formulation of evaluation criteria and weight allocation. In addition, AI technology can also provide richer information input for design evaluation by processing multimodal data (such as text, images, etc.) (Liu & Kim, 2023). However, although AI technology has shown great potential in design evaluation, it is still necessary to pay attention to the importance of human-computer collaboration. Design evaluation is not just a data processing problem, but also involves creativity, insight, and judgment of humanistic values. Therefore, future design evaluation systems should strive to combine the data processing and analysis capabilities of AI technology with the creativity and insight of human designers to build a more comprehensive and efficient evaluation framework (Liu & Kim, 2023; Rodgers & Huxor, 1998).

In general, design evaluation is a key link in the product development process. Its significance lies in its ability to effectively guide design decisions and improve the efficiency and success rate of product development. However, it also faces many challenges, such as the determination of evaluation standards, information uncertainty, and coordination of interests of multiple parties. With the development of technology, especially the application of AI technology, we are expected to develop more intelligent and accurate design

evaluation methods, thereby promoting product design and innovation to a higher level. This requires not only technological breakthroughs, but also continuous exploration and improvement in practice to adapt to the ever-changing market demands and technological environment.

In the era of globalization, what conditions make the world livable has become a critical issue. Latour (2009) believes that this issue is significant for designers. It concerns the survival of billions of humans and trillions of other organisms and involves the combination of nature and society. Latour and Peter proposed two viewpoints: sphere theory and actor-network theory (Latour, 2009; Sloterdijk et al., 2017). These two theories seem different, but both aim to bridge the gap between nature and society. Sloterdijk et al. (2017) further pointed out that the globalization of electronic communications represents the third wave of globalization, marking that we have entered a new era of interaction. The advancement of machine learning technologies has enabled the successful integration of artificial intelligence into design practices in contemporary research. For example, the Luban system that Alibaba Intelligent Design Laboratory developed can intelligently generate advertisements based on user behavior and preferences, improving design efficiency (Yang et al., 2018).

Nowadays, deep learning is a popular field of AI research, and stochastic gradient descent is the train weight of neural networks (Lecun et al., 2015). Neuroevolutionary is also another popular artificial intelligence research area. This method uses evolutionary algorithms to optimize neural networks (Kim, 2016; Lecun et al., 2015). It can meet the growing computing potential. Ideas can provide rich resources for the inspiration and mixing of deep learning, reinforcement learning, and machine learning. There are branches in the field

of artificial intelligence. Studies have confused the problems of automation, fuzzy logic, machine learning, and deep learning. This situation may lead to a lack of understanding or consistency in the future of artificial intelligence. Therefore, it is necessary to define important concepts and boundaries to ensure that researchers can accurately estimate and measure the impact of AI from the perspective of experience and norms. Research has shed light on this area.

There are growing concerns that robots and artificial intelligence will replace professions. To remain relevant in the ever-changing industrial environment, future employees must identify opportunities, change the industry, and provide innovative solutions to meet the challenges of increasing globalization in the information age (Rampersad, 2020). It has been argued by Rampersad (2020) that in the age of artificial intelligence, essential skills include critical thinking, problem resolution, effective communication, and team cooperation. All those abilities have significant impacts on the development of innovation. McCardle (2002) also discussed issues related to integrating emerging technologies into future designer education. With the rapid development of technology, industrial designers and educators face the challenge of entirely using these advanced technologies. Those challenges are particularly severe in higher education, which usually encourages adopting new innovative technologies (McCardle, 2002). For example, McCardle's research (2002) describes a model used to introduce the field of artificial intelligence to industrial design undergraduates. The designer's future role requires understanding the standards for practical usability, functionality, and product semantics and understanding the user's abilities, confidence, and product perception (McCardle, 2002).

The application of artificial intelligence can provide innovative and exciting opportunities for the design of intelligent and interactive products. In addition to education, the application of artificial intelligence technology in the industry enjoys a similar developing trend. Alibaba Group used Artificial intelligence technology and multimedia content design, which has created one of China's most significant High-tech business ecosystems (Liu et al., 2019) . For example, Liu and his colleagues (2019) introduced the general processing flow of AI-assisted design tools in the Alibaba business ecosystem. Considering that artificial intelligence technology can play a huge role in producing multimedia content, Alibaba could combine multimedia content design with artificial intelligence technology to expand application scenarios, such as AI-assisted design, graphic design, video, and webpage generation (Liu et al., 2019).

The 21st century is a significant period of technological development, especially in the field of information technology. These technological advancements have fundamentally changed the paradigm of global manufacturing, ushering in an era characterized by intelligent production systems. Contemporary research suggests that adopting intelligent manufacturing is a strategic priority for countries aimed at ensuring industrial leadership (Zhou et al., 2018). This technological transformation has triggered a global response, with major economies implementing comprehensive strategies to welcome the Fourth Industrial Revolution. It is worth noting that Germany has taken the lead in promoting this movement through its Industry 4.0 initiative (Liu, 2015), while China's "Made in China 2025" policy framework explicitly identifies intelligent manufacturing as a core strategic goal (Zhou et al., 2018). A key research focus in this field involves systematically integrating emerging information technologies into

traditional manufacturing processes. Scholars have identified the next generation of intelligent manufacturing as a synergistic combination of artificial intelligence and cutting-edge production methods (Zhou et al., 2018). In this constantly evolving environment, product design, as a fundamental component of manufacturing, needs to integrate AI based methods to maintain relevance and competitiveness.

In recent years, the application of artificial intelligence (AI) technology in product design has become a transformative trend, providing unprecedented opportunities to improve creativity, efficiency, and innovation (Quan et al., 2023). The fusion of big data and AI algorithms is revolutionizing product design methods, enabling designers to process and analyze massive amounts of information to optimize design outcomes (Zhang et al., 2022). This technological shift is not just an incremental change, but a change in basic assumptions in the way products are conceptualized, developed, and evaluated (Liu, 2015).

Researchers must address the importance of AI-assisted product design evaluation. The study pointed out that understanding consumer reactions and preferences for AI-designed products is crucial to developing effective evaluation models (Yang et al., 2021). AI-driven assessment systems have the potential to significantly improve the decision-making process, improve product quality, and accelerate time to market by providing objective, data-driven insights throughout the design lifecycle (Chong et al., 2022; Verganti et al., 2020). Existing research focuses on the application of AI in the product design generation stage, while the discussion on the design evaluation stage is limited. Liu's research (2015) shows that applying AI algorithms in cognitive thinking storage and feedback mechanisms can

effectively promote product design optimization. Yang et al. (2021) further proposed an intelligent product morphology design method based on cognitive dynamics and spider web structure, emphasizing the importance of selecting feasible solutions from intelligent design sketches. These studies have laid the foundation for the application of AI in design evaluation, but more systematic and comprehensive exploration is still needed.

In the field of product design evaluation, the application of AI technology can be divided into the following three levels. AI assisted evaluation refers to the use of AI as an auxiliary tool to provide data support and preliminary evaluation recommendations to designers, but the final decision-making power still lies in the hands of human designers. This model emphasizes human-machine collaboration, with AI playing a role in enhancing rather than replacing human designers. The evaluation of product design driven by artificial intelligence focuses on integrating artificial intelligence into the entire evaluation workflow, including multidimensional analysis of functional performance, aesthetic quality, and user interaction indicators. This approach emphasizes the systematic and comprehensive evaluation. AI driven evaluation represents AI playing a leading role in the evaluation process, independently completing evaluation tasks through technologies such as deep learning. This model emphasizes the autonomy of AI, but its application in subjective design evaluation is currently limited.

Based on the objectives of this study and the concept of human-machine collaboration, this thesis will uniformly adopt the term "AI assisted product design evaluation". This strategic approach emphasizes the complementary role of artificial intelligence in improving the efficiency and accuracy of the

design evaluation process, while affirming the indispensable leadership role of human designers in creative decision-making.

The application potential of AI in product design evaluation is enormous. First, AI can quickly process and analyze copious amounts of design data, extract valuable information, and provide an objective basis for design decisions. This ability is particularly suitable for handling complex multi-dimensional design problems. As Quan et al. (2023) emphasized, AI can consider multiple aspects, such as functionality, aesthetics, and user experience, at the same time to provide comprehensive evaluation results. The AI system can continuously optimize the evaluation model through machine learning. With the accumulation of data, the accuracy and reliability of the evaluation will continue to improve. This kind of self-learning and adaptability is not available in traditional assessment methods.

AI-assisted evaluation can also significantly improve design efficiency. Designers can get feedback faster by automating the evaluation process, resulting in faster iterations and shorter product development cycles. Verganti et al. (2020) pointed out that artificial intelligence can create continuously updated user-centered solutions. AI significantly impacts the effectiveness of the evaluation model by learning iteratively. Real-time data and user feedback can guide design decisions, which allows more personalized and innovative products to be designed and produced.

However, the application of AI in product design evaluation also faces challenges. The first issue is data quality and availability. The performance of AI models depends on the quality and quantity of training

data. In product design, obtaining a large amount of high-quality evaluation data can be challenging. Secondly, the interpretability of AI models is also a critical issue. Design evaluation usually involves subjective factors. Making AI evaluation results transparent and understandable to designers and stakeholders is a key issue. More research is needed in the future to address those issues. In addition, constructing a human-machine collaboration model is also a principal issue. How to fully utilize AI data processing and analysis capabilities while maintaining the creativity and insight of human designers is a direction that future design practices need to pursue. Chong et al. (2022) emphasized that human designers' effective use of AI input is crucial to success, and human ability and confidence play a key role in collaborating with AI systems.

The significance of this thesis is reflected in the following aspects.

The importance of design evaluation in the product development process is explored in depth. Effective design evaluation can significantly reduce the risk of product development and improve resource utilization efficiency. The nature of design evaluation as a complex multi-criteria decision-making problem is clarified. This study emphasizes that design evaluation involves multiple aspects, including task-related factors and decision-related factors, which helps to understand the complexity of design evaluation more comprehensively.

- The challenges faced by design evaluation, such as information uncertainty, imprecision of decision makers' judgments, and complexity of group decision-making, are explored. The identification of these

challenges provides directions for future research. The key role of design evaluation in the product design process is emphasized. Good evaluation can avoid expensive redesign costs in the later stage, which is crucial for the competitiveness of enterprises.

- The complexity of the modern product design decision-making process is pointed out, involving multiple stakeholders, such as engineers, consumers, designers, and enterprise managers. This multi-party participation feature increases the difficulty of decision-making and emphasizes the necessity of effective evaluation methods.
- Innovative ideas for design evaluation methods are proposed, including artificial intelligence, experimental evaluation, and online evaluation. The comprehensive application of these methods can provide more comprehensive and accurate evaluation results for product design. It also provides case references for future design education.
- The thesis discusses the potential of artificial intelligence and computer technology in simplifying the evaluation process and improving evaluation efficiency. This points out the direction for the development of future design evaluation methods. It emphasizes that design evaluation should not only focus on technical factors, but also consider multiple factors such as cost and user experience, which will help develop a more comprehensive and flexible evaluation method.

This study provides comprehensive theoretical and practical guidance for this emerging field by systematically exploring, implementing, and validating

AI-driven product design evaluation models. The research results have a meaningful impact on product design theory and practice, promote the application of AI technology in the creative field, and provide direction for future design methods and tools. By deeply understanding the potential and limitations of AI in design evaluation, this study will contribute to building a more efficient and innovative design ecosystem and promote product innovation and user experience.

1.2 Theoretical Foundation

The main objective of this study is to develop and validate an AI-assisted product design evaluation system. The system aims to bridge the gap between traditional design practices and innovative AI technology, providing designers with an evaluation tool to assist them in refining concepts. The AI product design evaluation system proposed in this study aims to provide comprehensive and objective evaluation results by adopting deep learning algorithms and multimodal data analysis. The study's theoretical basis includes the following aspects: artificial intelligence theory, product design theory, and human-computer collaboration theory.

1.2.1 Artificial Intelligence Theory

Artificial intelligence (AI) theory is one of the core foundations of this study. It includes related theories such as machine learning, deep learning, and natural language processing, which provide a basis for the construction of AI models. Machine learning theory, especially deep learning, enables AI systems to learn complex patterns and features from substantial amounts of data (Lecun et al., 2015). This deep learning ability is crucial for processing multi-dimensional, unstructured data in product design evaluation.

Natural language processing (NLP) theory provides theoretical support for the system to understand and analyze text data such as design descriptions and user feedback (Siddharth et al., 2022). Through NLP technology, the system can extract key information from the text, understand the design intent, and perform semantic analysis, thereby achieving more intelligent evaluation.

Computer vision theory provides a theoretical basis for the system to analyze the visual features of products (Chen et al., 2020). This enables the system to extract visual features such as shape, color, and texture from product images or 3D models for aesthetic evaluation and functional analysis.

The development of multimodal classification algorithms has also benefited from the progress of deep learning technology. By combining information from different modalities of image and text data, these algorithms can better understand and analyze complex design concepts (Cai, 2024; Radu et al., 2018; Ran et al., 2022; Zhou et al., 2018; Zhu et al., 2020). This theory is significant for improving production efficiency, resource utilization, and product quality.

1.2.2 Product Design Theory

Product design theory provides evaluation criteria and a framework for this study. Design thinking theory emphasizes a human-centered, iterative optimization design process, which is highly consistent with the goals of AI-aided evaluation systems (Brown, 2008). User-centered design theory emphasizes that products should meet user needs and expectations, which

provides important guidance for the evaluation criteria of AI systems (Norman, 2013).

In addition, product life cycle theory and sustainable design theory also provide important theoretical perspectives for the evaluation system, enabling the system to evaluate product design from a longer-term and more comprehensive perspective (Ceschin & Gaziulusoy, 2016).

1.2.3 Human-computer Collaboration Theory

The human-computer collaboration theory is an important theoretical basis for this study. This theory emphasizes that artificial intelligence (AI) is an auxiliary tool for designers rather than completely replacing human designers (Peuter et al., 2023) . In the product design process, the human-computer collaboration model can give full play to AI's data processing and analysis capabilities while retaining the creativity and insight of human designers.

The theory of human-computer collaboration believes that current AI technology has yet to separate from human intelligence completely, and its design and application still require human intervention (Feuston & Brubaker, 2021). In product design, AI is a collection of past achievements from human product design. However, human designers' creativity, originality, and open-mindedness are still crucial. This collaborative model enables design thinking to stand on the shoulders of AI while eliminating tedious literature and theoretical research, thereby promoting innovation in design thinking.

The theory of human-computer collaboration also emphasizes that AI technology can replace and assist human designers in completing specific product design tasks (Tsang & Lee, 2022) . Although AI has not entirely changed the entire product design process, it has brought innovations in product design methods. For example, AI can interact with designers to provide survey results of similar designs in the past and conduct more in-depth analysis and judgment on the success of these designs during the product research stage (Feuston & Brubaker, 2021).

In addition, human-computer collaboration theory also emphasizes that AI-assisted tools should have multiple interaction modes to adapt to distinctive design problems (Peuter et al., 2023) . To provide the best assistance, AI assistants should have multiple interactions, which should help designers to the greatest extent and provide information to AI assistants (Peuter et al., 2023) . This variety of interaction modes helps to enhance the effect of human-computer collaboration. Human-computer collaboration theory also states that to achieve effective collaboration, AI assistants need an accurate user model to understand the designer's behavior (Amershi et al., 2019; Peuter et al., 2023) . Although the theory is not limited to a single modeling paradigm, computational rationality is considered a particularly promising option that can produce scientifically sound models (Peuter et al., 2023) . Finally, the theory of human-computer collaboration emphasizes the development trend from human-computer interaction to deep coordination, and the importance of trust building as a breakthrough (Amershi et al., 2019; Feuston & Brubaker, 2021). With the rapid development of AI technology and its application in specific groups of people, intelligent agents begin to have clear interpersonal relationships, shift from passive to active interaction, and have emotional judgment and feedback intelligence (Amershi et al., 2019) .

This development trend further strengthens the importance and potential of human-computer collaboration.

In summary, the human-computer collaboration theory provides an important theoretical basis for this study, guiding us to fully utilize the advantages of AI technology in product design while maintaining the creativity and insight of human designers to achieve the best results of human-computer collaboration.

The organic combination of these theories provides a solid theoretical basis for developing an AI-assisted product design evaluation system. By integrating advanced theories in multiple fields, such as artificial intelligence and human-computer collaboration, this study aims to develop an intelligent evaluation system with both evaluation capabilities and the ability to collaborate with designers. In general, the theoretical basis of this study reflects the integration of contemporary design theory and artificial intelligence technology. This study explores the integration of artificial intelligence and design evaluation methods, aiming to change the design evaluation methods, improve the efficiency of the design process, and contribute new ideas and teaching tools to the progress of design education.

1.3 Research Objectives

The existing research on artificial intelligence enhanced design has identified significant research gaps in how artificial intelligence can optimize product development workflows. Based on these findings, this survey aims to develop an intelligent design support system and systematically evaluate its impact on the design process. The main research objectives involve the development and

empirical verification of an AI based design evaluation framework, with the following three specific objectives.

Objective 1: Optimize product design evaluation system using artificial intelligence technology

This objective aims to explore how to integrate artificial intelligence technology into the product design evaluation process. This thesis will study how to use the data processing and analysis capabilities of artificial intelligence while retaining the creativity and insight of human designers. This system will be able to process multimodal data to provide a more comprehensive perspective for design evaluation.

Objective 2: Verify the effectiveness of AI-assisted design evaluation systems

This objective involves evaluating the performance of AI models in design evaluation tasks. Through comparative experiments and practical application research, this thesis will explore how AI technology can enhance the capabilities of designers and improve innovation, accuracy, and efficiency. This will help understand the practical application value of AI in product design evaluation.

Objective 3: Construct an AI-assisted product design evaluation framework

Based on the research results, this thesis will establish a comprehensive research framework to explain how AI technology can improve design evaluation methods. This framework will combine traditional and modern AI technologies to provide a more comprehensive solution for the field of

engineering design. The goal of this study is to present the application field of AI-assisted design to designers or researchers and provide directions for future research.

By achieving these research objectives, this study aims to fully explore the application potential of AI in product design evaluation and provide designers with an evaluation tool to assist in improving conceptual design. This study will also provide theoretical and practical evidence for the application of AI in the design field, optimize the product design innovation process, and improve the efficiency of product design and development.

1.4 Overview of the Thesis Structure

The thesis structure is shown in Figure 1.1.

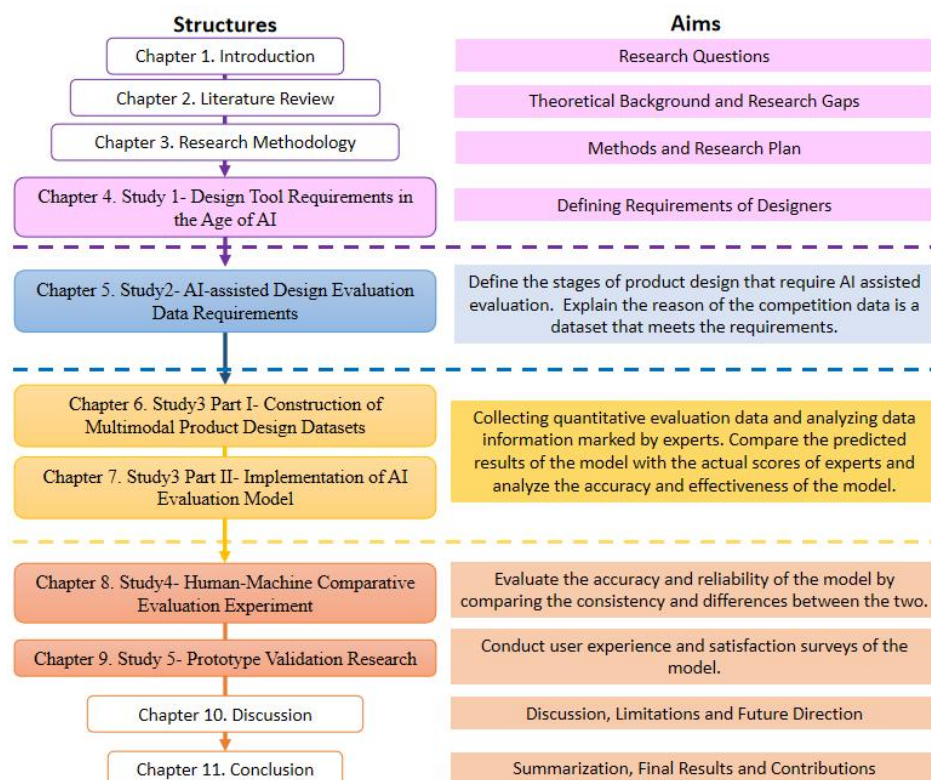


Figure 1.1 The Thesis Structure

Chapter 1 introduces the research background and motivation, clearly states the research objectives, and defines the research scope. This chapter also discusses the significance and value of the research, outlines the basic framework and direction of the entire research for readers, and lays the foundation for a deeper understanding of the research content.

Chapter 2 systematically reviews the previous research related to the application of artificial intelligence in design evaluation and discusses the current application status of AI technology in product design and its challenges. This chapter also explores the gaps in theory and practice in current research and identifies research gaps and questions, providing theoretical support and direction for the following research.

Chapter 3 introduces the methodological framework adopted in this study in detail, including research design, data collection methods, and data analysis techniques. At the same time, this chapter also elaborates on the specific steps of the research process and experimental design to ensure the scientific and operability of the research.

Chapter 4 analyzes designers' basic needs and expectations when using AI-assisted design tools and determines the functions and performance indicators that AI tools should meet. Through demand analysis, this chapter provides specific guidance and reference for developing AI tools.

Chapter 5 defines the specific links in the product design stage that require AI-assisted evaluation and explains how the data in these links can meet competitive needs. At the same time, this chapter clarifies the specific requirements that the data set should meet and provides standards for data collection and processing.

Chapter 6 details how to collect and construct a multimodal product design dataset for AI evaluation. This chapter discusses the data types, collection methods, and preprocessing procedures, providing a detailed operational guide for constructing the dataset.

Chapter 7 introduces the construction process of the AI evaluation model, including the design, training, and optimization strategies, and demonstrates the application of the model in actual design evaluation. Through empirical research, the effectiveness and practicality of the model are verified.

Chapter 8 evaluates the accuracy and reliability of the AI model through human-machine comparison experiments and compares the consistency and difference between the AI evaluation results and the expert scores. The experimental results in this chapter provide a basis for optimizing and improving the AI model.

Chapter 9 conducted a user experience and satisfaction survey to verify the practicality and user acceptance of the AI assessment tool. By collecting user feedback, this chapter provided valuable user insights for further product optimization.

Chapter 10 provides an in-depth discussion of the findings, analyzing the study's limitations and directions for future research. Through critical analysis, this chapter proposes future research directions that address the issues discovered.

Chapter 11 summarizes the main results and contributions of the research and explores the practical application prospects and promotion value of AI evaluation models in the design field. This chapter puts forward the conclusion of the research, providing reference and inspiration for research and practice in related fields.

This chapter outlines the research background, significance, and overall framework of AI-assisted product design evaluation, laying the foundation for the following research. By elaborating on the theoretical basis and structure of the research, this study emphasizes the importance of exploring the application of AI in product design evaluation. This provides a clear direction for the literature review in the next chapter, guiding the thesis to explore the current research status in depth, finding research gaps, and thus clarifying this study's specific problems and aims. Although this chapter provides an overall framework for the research, a literature review is still needed to determine the research gaps and directions.

Chapter 2. Literature Review

This chapter will comprehensively review existing AI-assisted product design evaluation research and find research gaps. The application status, challenges, and opportunities of AI in product design evaluation will be explored by analyzing the current research status. This analysis will help the thesis clarify research questions and aims and provide theoretical support and guidance for the following research. The literature review in this chapter will be discussed in parts according to the research logic. Finally, the research gaps and questions will be drawn, laying the foundation for the next empirical research.

2.1 Current Status of AI Application in Product Design Evaluation

The development of artificial intelligence has made AI driven design evaluation a key factor in improving creative innovation and operational efficiency in the product development process. AI technology is redefining all stages of product design by providing advanced data analysis, automated design processes, and tools to enhance creativity. This section discusses the status of AI applications in product design evaluation through a comprehensive analysis of relevant literature, including its methodology, benefits, challenges, and future development prospects.

2.1.1 Current Status and Impact of AI in Product Design

Artificial intelligence is vital in improving design efficiency and stimulating innovation potential. The application of AI covers all stages of the design

process, from initial conceptualization to final product evaluation. For example, the AI-assisted design (AIAD) method mentioned by Quan et al. (2019) can effectively explore and optimize various design solutions using advanced search techniques and algorithms. In complex design tasks such as sustainable urban development projects, AIAD can quickly analyze multiple possible urban layouts and building plans, considering environmental impacts, cost-effectiveness, and social factors, thus determining the best design solution.

In addition, Manavis (2023) emphasized the advantages of combining AI with existing technology tools, pointing out that this combination significantly improves the efficiency of product and packaging development processes. AI frees up designers' time by automating repetitive tasks in design, such as data collection and preliminary analysis, allowing them to focus on more creative and strategic design activities. In industries where rapid prototyping and iterative design are critical, such as consumer electronics and automotive manufacturing, the application of AI can significantly shorten the time it takes for products to go from concept to market and make dynamic adjustments based on real-time user feedback (Yan & Li, 2020; J. Zhou et al., 2018). AI technology also plays a significant role in all stages of product design, including inspiration acquisition, creativity and concept generation, concept evaluation, optimization, decision-making, and modeling (Yüksel et al., 2023). Especially in the creative generation stage, AI can help designers overcome design rigidity and promote the generation of more innovative and breakthrough ideas.

AI tools, especially language processing models such as ChatGPT, have begun to change design thinking and decision-making processes profoundly.

These tools help designers make more scientific and reasonable decisions in the design process by providing data-based insights and predictions. Al-Sa'di and Miller (2023) showed that the application of AI in design thinking can improve the quality of decision-making and promote innovative solutions by discovering hidden design opportunities. In the product design process, AI technology has significantly changed how products are conceptualized, developed, and optimized (Liu & Kim, 2023). Through a careful evaluation of the effects of AI technology application, research shows that AI technology has had a profound impact at all stages of product design (Liu & Kim, 2023). AI technology not only enhances the capabilities of designers but also changes the structure and flow of the entire design process. Integrating AI simplifies the processing of complex data, enabling designers to optimize product design by leveraging historical data, market trends, and consumer behavior analysis. This data-driven design approach shows exciting potential in improving product success and market responsiveness. However, despite considerable progress in the application of AI in design thinking, more than research on its application in dealing with unclearly defined user needs and expanding into new markets is required. This research gap suggests that future research needs to explore how to integrate AI better to solve these complex problems.

AI plays a key role in improving the multidimensionality of design evaluation, especially in evaluating the aesthetics and functionality of designs. Botros et al. (2023) explored the application of AI in architectural aesthetic evaluation. By analyzing design examples and user feedback, AI was able to identify which design elements resonate most with users and how to improve the design to increase user satisfaction. These insights are critical to ensuring that design solutions are technically feasible and resonate with users aesthetically

and emotionally. In addition, with the development of virtual reality (VR) and augmented reality (AR) technology, the application of AI on these platforms provides new evaluation tools for product design (Fan et al., 2022). Designers can quickly prototype and test modern design concepts in a virtual environment, obtain instant user feedback, and adjust the design based on this feedback. The immediacy and interactivity of this design evaluation enhance the flexibility and user-centricity of the design process.

In practical applications, AI technology has been used to analyze user needs and design standards, especially in big data environments. For example, Wang et al. used AI algorithms to optimize product appearance and color matching based on the diverse needs of users and the functional requirements of products (Liu & Kim, 2023). By establishing a support vector machine (SVM) appearance color intention model, designers can more accurately capture user preferences and transform them into specific design elements (Liu & Kim, 2023).

2.1.2 Challenges and Future Prospects of AI Applications

Although AI technology has brought benefits to product design evaluation, it also faces challenges in its widespread application. As intelligent elements are increasingly integrated into product design, the application of AI technology in product design has become a mainstream direction. It has shown great application prospects in fresh fields, extending to human-computer interaction services (Liu & Kim, 2023). This trend has not only changed the way designers work but is also reshaping the interaction mode between users and products. Among them, the interpretability of AI systems is a critical issue. As AI models become increasingly complex, the transparency and

interpretability of their decision-making process become lower and lower, which is particularly problematic for occasions where design decisions need to be accurately explained. Zhu et al. (2021) pointed out that this problem is particularly prominent in human-computer interaction (HCI) environments, which may lead to reduced user acceptance of AI-driven design solutions.

The potential bias of AI algorithms is also an issue that cannot be ignored. These biases may arise from imbalanced or insufficient training data, resulting in AI-output design solutions that fairly serve only user groups. Morley et al. (2019) emphasized that it is crucial to consider ethical issues during AI development, which helps identify and mitigate these biases and ensure the fairness and universality of design solutions. At the same time, the interpretability of AI systems and the need for domain-specific knowledge remain important challenges (Liu & Kim, 2023). Designers and engineers must understand the logic behind AI-generated design suggestions to make informed decisions. Maintaining a balance between algorithm optimization and human creativity is also an issue that needs continued exploration.

Looking ahead, the application prospects of AI in product design evaluation are very promising. With the further development of technologies such as machine learning and neural networks, the application of AI in design tools will become deeper and more extensive. The parametric design and generative design methods discussed by Hegazy and Saleh (2023) show that AI can help designers explore abstract design concepts and automatically generate multiple design options based on specific design parameters. This ability will improve the quality and efficiency of design, making products more in line with market and user needs.

The application of deep neural networks (DNNs) in product design has shown exciting potential. For example, by using AI techniques such as particle swarm optimization and genetic algorithms, designers can optimize product design parameters more effectively (Liu & Kim, 2023). These techniques improve design efficiency and produce more innovative and user-friendly design solutions. Actual case studies have also confirmed the effectiveness of AI technology in product design. Taking seat design as an example, through the application of AI technology, the study found that users prefer leather materials and diverse materials while giving lower scores to painted wood (Liu & Kim, 2023). Regarding product structure and shape, users prefer arc and curved structures rather than straight structures. These findings provide valuable guidance for designers to help them create products that better meet user preferences.

New research directions are emerging in integrating artificial intelligence and product design. However, to fully realize the potential of AI in product design evaluation, multiple challenges need to be addressed, including technical, ethical, and operational challenges. A key issue is how to leverage the capabilities of AI while maintaining human creativity effectively. This integration can enhance product innovation and efficiency and improve the decision-making process and optimization of design parameters.

In general, the application of AI in product design evaluation is undergoing a rapid evolution and modernization. By integrating multimodal data, deep learning techniques, and traditional design methods, AI is bringing revolutionary changes to the field of product design. However, to fully realize

the potential of AI, it is necessary to find a balance between technological development, ethical considerations, and practical applications. Future research should focus on developing more intelligent, explainable, and user-friendly AI design tools while exploring how to maximize the constructive collaboration between AI and human creativity. These limitations will be discussed in detail in the next section.

2.2 Limitations of Existing Research

Despite the promising application of AI in product design evaluation, significant limitations in existing research hinder its effective implementation. Through a comprehensive analysis of relevant literature, this study summarizes the limitations in the following significant aspects: insufficient exploration of AI capabilities, insufficient data quality, and challenges in integrating AI into existing design processes. This section will explore the current barriers and challenges of applying AI in product design evaluation.

2.2.1 Limitations of Design Evaluation Studies

Design evaluation studies are essential for assessing the effectiveness and efficiency of various design interventions across multiple domains. However, these studies face limitations that may hinder their effectiveness and applicability. This section synthesizes the existing literature and explores the challenges inherent in design evaluation studies, focusing on stakeholder involvement and the complexity of real-world applications.

Stakeholder engagement poses a significant challenge in design-evaluation research. Instructional designers have been found to often prioritize informal evaluation methods over formal evaluation methods, which may not

adequately capture the perspectives of all relevant stakeholders (Williams et al., 2011). This disconnect can result in evaluations that do not fully represent the needs and expectations of end users, resulting in designs that do not achieve their intended effects. Additionally, DeV Vaughn and Stefaniak (2020) highlight that a lack of stakeholder interest in measuring the effectiveness of training programs can further complicate the evaluation process as it can lead to inadequate data collection and analysis. This challenge is particularly evident in settings where evaluation is less valued, such as in government and military settings, where resources are often allocated to new instructional materials rather than evaluating existing ones (DeV Vaughn & Stefaniak, 2020).

Resource allocation is another key factor that limits the effectiveness of design-evaluation research. Researchers face constraints related to funding, staffing, and time, which can significantly impact the scope and depth of their evaluations. point out that peer-reviewed journals are often limited in their flexibility to accept different evaluation designs, which may prevent researchers from pursuing innovative evaluation methods (Celinska, 2021). In addition, the challenges of conducting evaluations in resource-limited settings may result in underpowered studies that fail to produce meaningful results, as they emphasize in their discussion of regression discontinuity designs (Louie et al., 2016). Lack of adequate resources can also hinder the conduct of comprehensive evaluations that consider a variety of dimensions, such as process, economic, and equity measures, which are critical to understanding the broader impact of design interventions (Jones et al., 2014).

The complexity of real-world applications further complicates design evaluation research. For example, point out that in healthcare IT design, the disconnect between designers and end users often results in evaluations that

do not accurately reflect users' interactions with technology (Kulp & Sarcevic, 2018). This gap may lead to designs that are inconsistent with user needs, reducing their effectiveness. In addition, challenges associated with implementing evaluations in dynamic environments (e.g., education) can result in significant variation in intervention implementation and outcomes, making it difficult to draw clear conclusions from evaluation studies (Kane et al., 2012). Recognizing that interventions must be tailored to specific contexts to achieve desired outcomes underscores the need for adaptability in evaluation design (Cornet et al., 2020).

In addition, the subjective nature of design evaluation presents another layer of complexity. As Thurston noted, different stakeholders often have different perceptions of the value of a design, which leads to challenges in establishing standardized evaluation criteria (Selva et al., 2014). This subjectivity can lead to conflicting interpretations of evaluation results, further complicating the decision-making process regarding design improvements. Furthermore, as discussed, the lack of uniform evaluation metrics can hinder the ability to quantitatively assess design effects, particularly in areas where sensory aspects are critical (Kato & Tsuda, 2018). The lack of clear metrics can lead to ambiguity in assessing the success of design interventions, impacting their implementation and sustainability.

Design evaluation research is fraught with limitations that can hinder its effectiveness and applicability. Methodological limitations, stakeholder engagement challenges, and the complexity of real-world applications all contribute to the difficulties faced by researchers in this field. To improve the effectiveness and impact of design evaluation, researchers must adopt more flexible and inclusive methods that consider the diverse perspectives of

stakeholders, allocate adequate resources, and account for the complexity of real-world settings. By addressing these limitations, the field of design evaluation can evolve to better meet the needs of practitioners and end-users.

Existing research on the exploration of AI tools in product design still needs to be in-depth and comprehensive enough. Al-Sa'di and Miller (2023) pointed out that although AI language models such as ChatGPT have potential in the design stage of technological innovation, their application research in solving undefined user needs and remote markets is still insufficient. This reflects that the integration of AI and design thinking principles has yet to be fully realized, and future research is needed to explore this intersection further. Kong (2020) showed that AI is rarely used in creative fields such as art and design. This finding highlights the need for interdisciplinary research to explore how AI can be effectively applied in various design fields. Future research must prioritize applying AI to areas of creativity and artistic expression.

2.2.2 Challenges of Data Quality, AI Explainability, and Evaluation Framework

Data quality and availability significantly affect the practical application of AI in product design evaluation. Vadla (2024) discussed the application of sentiment analysis in product design evaluation, noting that reliance on customer feedback data may lead to biased or unrepresentative results. Ahmad et al. (2022) further pointed out that a significant obstacle manufacturing companies face in implementing AI is the need for more data analysis talent and skills. This affects the ability to effectively use AI tools and limits organizations from fully realizing the potential of AI in product design

evaluation. These findings highlight the importance of improving data quality and enhancing workforce data analysis skills, while also pointing out the need for educational institutions and industry stakeholders to pay attention to this systemic problem together.

Existing research needs a comprehensive framework to guide the integration of AI in the product design process. Jylkäs et al. (2019) emphasized the need for systematic research on how service design can promote the scalability of AI-enabled systems. Organizational challenges in AI technology assimilation were systematically documented by Bérubé et al. (2021) using the Delphi method, with particular emphasis on data handling limitations and workforce skill gaps specific to artificial intelligence. In addition, Kong (2024) pointed out that the lack of standardized evaluation protocols to evaluate the application of AI in product design may make it difficult to obtain reliable evidence on the effectiveness of AI. These findings highlight the importance of establishing a standardized evaluation framework to improve the credibility and reliability of AI applications in product design.

Herskovit (2021) noted that AI models have difficulties with interpretability, which can cause users to doubt their reliability and validity. In product design, stakeholders must understand the reasoning behind AI-generated evaluations to trust and adopt these technologies. The opacity of AI's decision-making process may significantly hinder its acceptance in the design evaluation environment. Khanolkar et al. (2021) further argued that while AI can assist engineering design, a fully automated design process remains unattainable, emphasizing the importance of maintaining a collaborative approach that leverages the combination of AI capabilities and human creativity.

These limitations highlight the shortcomings of existing research on the application of AI in product design evaluation and point out the direction for future research. To overcome these challenges, developing more comprehensive and effective methods to integrate AI technology with product design evaluation is necessary. In this regard, multimodal data and deep learning technology provide new possibilities for improving the application effect of AI in product design evaluation. The following section will explore in detail the potential of multimodal data and deep learning in product design evaluation, as well as how to use these advanced technologies to overcome the limitations of current research.

2.3 Potential of Multimodal Data and Deep Learning in Product Design Evaluation

With the rapid development of technology, multimodal data, and deep learning have shown great application potential in product design evaluation. This section will focus on how these technologies can enhance the design process, improve user experience, and promote collaboration among design teams.

2.3.1 Improving the Comprehensiveness and Accuracy of Design Evaluation

The comprehensive use of multimodal data can significantly improve the comprehensiveness and accuracy of product design evaluation. Hu (2022) showed that the fusion of multimodal sensor data can enhance market adaptability and responsiveness to user needs. Said et al. (Said et al., 2017) emphasized the effectiveness of deep learning technology in processing

EEG-EMG data compression and classification. This technology is of excellent value in understanding various user interaction methods and can provide data support for design decisions and iterations.

Evaluation methods for deep learning models have become increasingly sophisticated, focusing on performance metrics, including accuracy, robustness, and ethical considerations (Boursalie et al., 2022). Such a comprehensive evaluation framework is critical to ensuring the developed models are compelling, trustworthy, and applicable in real-world scenarios. Design evaluation means evaluating a product's aesthetic and functional qualities through a multimodal perspective, combining user feedback and performance data, and iteratively improving the design.

2.3.2 Promote Product Innovation and Meeting User Needs

By analyzing multimodal data, designers can explore user preferences and behavior patterns in more detail. Yu et al. (2022) demonstrated how large-scale commercial multimodal representation learning can provide deep insights into user-product interactions, which is essential for customizing products to meet the needs of specific user groups. At the same time, Ayık (2023) showed that creative evaluation from a multimodal perspective can inspire more prosperous and innovative design ideas.

2.3.3 Implementation of Real-time Evaluation and Iterative Optimization

The combination of multimodal data and deep learning offers new possibilities for real-time product performance monitoring and evaluation. Bakalos et al. (2019) discuss methods for using multimodal data fusion and adaptive deep learning in monitoring critical systems. These methods are also

applicable to product design, enabling designers to quickly adjust and optimize designs based on real-time feedback. Deep learning methods have been used in product design to evaluate and optimize product aesthetics. Using image datasets and corresponding user ratings, researchers have developed convolutional neural networks (CNNs) that can evaluate the aesthetic appeal of product designs. This approach simplifies the design evaluation process and provides designers with actionable insights based on user preferences, thereby promoting a more user-centric design approach (Zhou et al., 2021). Incorporating multimodal data such as visual and textual feedback further enriches this evaluation process, allowing for a more comprehensive understanding of user perception.

2.3.4 Enhance Interdisciplinary Collaboration and Decision Support

Integrating multimodal data provides product design teams with a more affluent information base, which helps promote interdisciplinary collaboration and decision support. By processing and analyzing multimodal data with deep learning technology, design teams can obtain more comprehensive and insightful information about user needs and market trends (Yüksel et al., 2023). This information can be shared among team members with different professional backgrounds to promote effective communication and collaboration. For example, by combining image processing and natural language processing technology, designers can simultaneously analyze product appearance and user reviews to make more comprehensive and accurate design decisions (Liu & Kim, 2023).

Zheng (2024) introduced a variational neural network optimized using the Osprey optimization algorithm to evaluate product appearance aesthetics.

This algorithm improved accuracy, precision, and specificity significantly compared to the state-of-the-art. Integrating deep learning in this context allows for a more nuanced understanding of aesthetic values in industrial design, demonstrating how machine learning can improve product aesthetic assessment (Zhang, 2024).

Although multimodal data and deep learning have shown enormous potential in product design evaluation, there are still significant research gaps in practical applications. For example, the effective integration and alignment of data and how to ensure the interpretability and credibility of artificial intelligence evaluation are challenges that need to be addressed urgently. Future research should focus on developing standardized methodologies so that designers can effectively use these innovative technologies to ensure the optimization and innovation of the design process. Through the above discussion, the researcher can outline the application potential of multimodal data and deep learning in product design evaluation and propose new questions and directions for subsequent research to promote further development in this field.

2.4 Research Gaps and Research Questions

Although the application of AI in product design evaluation has shown significant potential, there are still important gaps in existing research that need to be filled by future research. This section will explore these research gaps and propose research questions for future academic exploration and technology development.

- **Clear definition of AI-assisted design requirements**

Although existing studies have explored the application requirements of AI in the design stage, an in-depth definition and analysis of the specific requirements and functions of AI tools at distinctive design stages still need to be provided. Existing literature emphasizes the importance of AI technology in product design. However, systematic guidance still needs to be on how to effectively integrate AI technology at various stages of the design process. More in-depth research is needed to clarify how AI technology can enhance human capabilities in terms of defining user needs, establishing success criteria, and building reasonable expectations.

- **Design and optimize evaluation data requirements and sampling strategies**

In AI model development, selecting appropriate samples and data types and processing these data to meet the accuracy and effectiveness of AI tools are issues that still need to be fully addressed in current research. Existing research emphasizes the importance of dataset type and breadth in meeting design requirements. However, how to create new datasets and reduce the interference of evaluators' subjective consciousness still needs further exploration. In addition, how to ensure that AI technology can effectively enhance human capabilities after completing fully automatic tasks is also a key issue.

- **Construction and application of multimodal design evaluation dataset**

Multimodal datasets play a key role in improving the performance of AI evaluation models, but how to effectively build and utilize these datasets and integrate different data modalities to optimize AI models remains a challenge. Existing research shows that multimodal data analysis can

reveal the complex interaction patterns between users and products, but how to transform these insights into specific design decisions still requires in-depth research.

- **AI evaluation model verification and user experience**

The methods and standards for verifying the effectiveness of AI evaluation models still need to be determined. At the same time, how to evaluate the user experience and satisfaction of these models in the actual design process also requires more research. Existing literature emphasizes the interpretability of AI systems and the need for specific domain knowledge, but how to effectively utilize the capabilities of AI while maintaining human creativity remains a significant challenge. In addition, how to maintain a balance between algorithm optimization and human creativity also needs further exploration.

This thesis proposes the following research questions based on the above research gaps.

- **RQ1: At which product design stage do designers need AI-assisted tools the most? What functions should these tools have?**

This question aims to clarify the application requirements of AI technology in each product design stage. By intensely studying the challenges and needs designers face at various stages, this study can better define the functions and characteristics of AI tools, thereby improving their practicality and effectiveness in the design process.

- **RQ2: What data types need to be collected and processed to develop AI-assisted design evaluation tools?**

This question focuses on the optimization of data requirements and processing strategies. By exploring the value and processing methods of diverse data types, this study can provide more comprehensive and effective training data for AI models, thereby improving the performance and applicability of the models.

- **RQ3: How do we build an AI product design evaluation model based on multimodal data to improve the accuracy and comprehensiveness of the evaluation?**

This question explores the application of multimodal data in AI design evaluation. This study can develop a more comprehensive and accurate evaluation model by studying how to effectively integrate and utilize data from different modalities (such as images and texts).

- **RQ4: How can the effectiveness of the AI-assisted product design evaluation model be evaluated? In practical applications, what are the similarities and differences between the model and the evaluation results of human experts?**

This question focuses on the validation and practical application of AI models. By comparing the evaluation results of AI models and human experts, this study can gain in-depth insights into the advantages and limitations of AI technology in product design evaluation and explore the best practices of human-machine collaboration.

These research questions aim to fill existing research gaps and provide theoretical and practical guidance for applying AI in product design evaluation. By answering these questions, this study will help promote the development of AI-assisted design evaluation tools and provide new methods and insights for design practice. At the same time, these questions also reflect the main challenges and opportunities of the current application of AI technology in the field of product design, including how to effectively integrate AI technology with traditional design methods, how to process and utilize multimodal data, and how to maximize the potential of AI technology while maintaining human creativity.

By systematically exploring these RQs, this thesis can advance theoretical research on AI in product design evaluation and provide specific guidance and tools for design practitioners, thereby promoting the widespread application and innovation of AI technology in product design. This is significant for improving design efficiency, enhancing product innovation capabilities, and better meeting user needs.

This chapter reviews existing AI-assisted product design evaluation research and shows key research gaps and directions. This study clarifies the research questions and aims by analyzing the existing literature, providing theoretical support for the following empirical research. This review provides an important basis for the design of the research method in Chapter 3 and helps us find a hybrid research method suitable for solving the research problem. However, the comprehensiveness of the literature review may be limited by the scope of the selected literature. The next chapter will design a specific research method based on the findings of this chapter.

Chapter 3. Research Methodology

This chapter elaborates on the methodological framework, data collection, and analysis methods used in this study. Based on the theoretical foundation and research gaps in the earlier two chapters, the thesis designs a mixed research method that includes qualitative and quantitative methods. This chapter will introduce five interrelated studies (Study 1-5), each addressing a specific research question (RQs) and adopting a corresponding research method. This multimethod, multi-stage research design aims to fully explore the possibilities of AI-assisted product design evaluation and ensure the reliability and validity of the research results.

3.1 Research Methods

This thesis adopts a mixed research approach, combining qualitative and quantitative methods, to comprehensively explore the complex issues of AI-assisted product design evaluation. This method, based on the multidimensional nature of the research question, provides deeper and more comprehensive insights. The use of mixed methods allows for an understanding of complex research problems from multiple perspectives and improves the reliability and effectiveness of research results through triangulation (Rath et al., 2022). Triangulation is a research method that enhances the reliability and effectiveness of findings by using multiple data sources, analytical methods, or theoretical perspectives. It is beneficial in mixed-nature fields such as information systems (IS) research, as it helps researchers understand complex problems from multiple perspectives. Triangulation can also help researchers systematically consider existing

theories without being driven by them, resulting in more useful, relevant, and up-to-date theoretical construction (Glaser et al., 2005).

The use of triangulation in research has the following advantages. It improves the credibility of research results by employing multiple methods and data sources. It allows for the validation of findings from different perspectives, enhancing the reliability of the results (Glaser et al., 2005). It reduces bias and assumptions, allowing researchers to systematically consider existing theories without being driven by them. It eases a comprehensive understanding of complex problems, especially in mixed-nature research fields like information systems. It simultaneously integrates qualitative and quantitative methods, such as considering aesthetic value and product functionality in product design research (Chumiran & Zainal Abidin, 2021).

When studying emerging technologies or phenomena, triangulation methods can effectively adapt to and explore these fields (Rath et al., 2022). In practical applications, triangulation can take various forms, such as combining analysis, testing, validation, and monitoring in the research of AI software quality assurance (Gezici & Tarhan, 2022) or using visual element research and principal solution research simultaneously in design cognition research to comprehensively understand product form (Chumiran & Zainal Abidin, 2021). By combining theoretical exploration and practical verification, triangulation can generate more relevant and up-to-date theoretical construction (Glaser et al., 2005).

Triangulation is a powerful research method particularly suitable for complex, interdisciplinary research problems, providing more comprehensive and

reliable results. This study proposes a theoretical framework for using AI technology to improve design evaluation methods, emphasizing AI's potential to enhance evaluation efficiency and accuracy, especially in quantifying design scheme evaluation results. This study involves multiple disciplines, such as computer science, design, and human-computer interaction, making it suitable for a mixed research method.

3.1.1 Qualitative Methods

In terms of qualitative research, this study adopted focus group interviews and grounded theory analysis methods to explore the real needs of designers for artificial intelligence tools.

3.1.1.1 Focus Group

Focus group interviews allow researchers to delve into designers' specific needs and expectations in the AI era, comprehensively covering research topics through structured discussion data (Gidumal et al., 2024). The choice of focus group interviews is based on the following key considerations. As a qualitative method widely used in social science research (Tremblay et al., 2010), focus groups can flexibly manage various design themes and fields. It allows researchers to interact directly with domain experts and obtain rich data (Denton & McDonagh, 2003; Peuter et al., 2023). This method is particularly suitable for exploring emerging and complex topics such as AI-assisted design. Because it can stimulate interaction among participants, generate collective intelligence, and delve into designers' real needs and concerns about AI-assisted tools (Peuter et al., 2023; Tremblay et al., 2010).

Among qualitative research methods, focus groups and interviews are the most common and effective methods for exploring well-defined topics. Each method has different advantages, depending on the nature of the research. In-depth interviews are ideal for gathering detailed insights into individual behaviors, thoughts, and perceptions. In contrast, focus groups are designed to obtain a wide range of opinions from participants who engage in structured discussions, allowing researchers to observe a range of perspectives on a particular topic (Martin & Hanington, 2012) . For the purposes of this study, the focus group method is a more appropriate method because the focus of this study is on AI-assisted product design, involving product designers as the main research subjects.

Focus groups can promote in-depth discussions on topics specifically selected by researchers, allowing for a dynamic exchange of ideas between participants. This interaction not only promotes the discovery of different perspectives, but also helps to synthesize these perspectives into coherent conclusions. Since the research revolves around clearly defined topics and target user groups, the focus group format (composed of a group of experts) enhances the relevance and depth of the data collected. Compared with individual interviews, focus groups provide a more collaborative environment, and collective insights can lead to more balanced and fair results(Martin & Hanington, 2012).

Furthermore, in a focus group setting, participants can provide inspiration to one another, thus enriching the discussion and leading to more reliable conclusions. Specifically, focus groups are suitable for in-depth research on complex issues aimed at predicting future trends and developments. This method involves structured conversations, and researchers can effectively

guide the discussion to ensure that each topic is thoroughly explored. This structured approach not only captures the collective expertise of the participants, but also closely aligns with the research goal of exploring the evolving landscape of product design influenced by artificial intelligence (Gordon et al., 1969) . Therefore, the using of focus groups is particularly suitable for achieving the goals of this study given their specific focus and the expert nature of the participants.

3.1.1.2 Grounded Theory

Grounded theory coding helps this study extract key concepts and theories from raw interview data, which is crucial for understanding the application of emerging technologies in the design field (Charmaz, 2015). Grounded theory is highly suitable for exploratory research (Charmaz, 2015) . Artificial intelligence and product design are emerging disciplines that have appeared with the development of computer technology and the industrial revolution. In the field of AI-driven product design, which lacks mature theories, grounded theory can help us construct a theoretical framework for AI-assisted design requirements based on designers' practical experience and perspectives (Peuter et al., 2023) . In addition, the importance of this study reflected in its emphasis on the viewpoint that AI should collaborate with designers rather than replace them (McCardle, 2002). AI-assisted design tools should support and use designers' creativity and problem-solving abilities rather than simply automating the design process (Figoli et al., 2022).

By combining focus groups and grounded theory, this study can by combining focus groups and grounded theory, this study can systematically collect and analyze data to understand how designers perceive the

relationship between artificial intelligence and the design process and their expectations and concerns about AI tools (Denton & McDonagh, 2003). This guides the development of AI-assisted tools that better meet the needs of designers (Peuter et al., 2023). The innovation of this method lies in its ability to be used not only for exploratory research but also for confirmatory research, such as verifying the practicality of design results through confirmatory focus groups (Sanabria-Z & Olivo, 2024). Therefore, this study was well suited to using grounded theory to analyze and explore qualitative data.

3.1.2 Quantitative Methods

This thesis uses quantitative research methods to explore the possibility and accuracy of AI-assisted product designers in evaluation tasks. This study uses questionnaires, natural language processing technology, computer vision, multimodal neural networks, and experiments in quantitative research.

3.1.2.1 Questionnaires

Questionnaire surveys collected a large amount of standardized data (Tubadji et al., 2021), which is crucial for understanding the everyday needs and attitudes of the designer community. In design research, questionnaire survey methods play a crucial role. Questionnaire surveys are regarded as the primary method for collecting data in various fields, including academic research, public institutions, and business organizations research and data collection work (Agrawal, 2010). It involves meticulous planning, construction, and writing of survey materials to elicit expected responses from participants (Agrawal, 2010; Taherdoost, 2022). In addition, questionnaire surveys are

quantitative and qualitative research methods to collect participant insights (Hendriana et al., 2023; Ohueri et al., 2019).

In the research of survey methods, questionnaire design is a key part that affects the quality and reliability of the collected data (Rattray & Jones, 2007). Researchers emphasize the importance of using pre-planned methods to determine the questionnaire's reliability and validity to ensure the collected data's robustness (Rattray & Jones, 2007). In addition, the questionnaire design should aim to reduce potential biases and errors that may affect the research results (Antoun et al., 2017). By following best practices in questionnaire design and development, researchers can improve the rigor and credibility of their research results.

During the questionnaire survey process for product design evaluation, researchers must consider a range of factors to ensure the effectiveness of the data collection process. Choosing survey design, constructing survey tools, and evaluating reliability and effectiveness are key stages of survey-based research (Kitchenham & Pfleeger, 2008). Researchers must also pay attention to the management of survey tools and next analysis of collected data to obtain meaningful insights (Kitchenham & Pfleeger, 2008). By following the systematic questionnaire survey design and data analysis methods,

researchers can obtain valuable information to provide insights for product design decisions. Questionnaire surveys provide a structured method for collecting feedback and insights when analyzing data obtained from expert product design evaluations. Using questionnaire surveys enables researchers to systematically collect data on experts' opinions, preferences, and criticisms of product design (Yang, 1989). The data collected through questionnaire surveys can be analyzed using various statistical methods to draw meaningful conclusions about the quality and efficacy of the evaluated product design (Cai, 2024).

In summary, the questionnaire survey method plays a significant role in design research, especially in evaluating product design through expert data analysis. By using carefully designed questionnaires, researchers can improve the quality of data collection and analysis in design studies. Through meticulous planning, structured survey material design, and rigorous data analysis, researchers can gain valuable insights, provide information for product design decisions, and contribute to progress in the field. Methodologically, using the questionnaire survey method and SPSS for statistical analysis is a scientific and effective research method. This method can collect the actual needs and perspectives of designers, which is in line with the emphasis in the literature on combining artificial intelligence technology with people's actual needs (Liu & Kim, 2023). A questionnaire

survey can help understand how designers view the application of AI technology in fresh design stages, which is of great significance for defining product design stages that require AI-aided evaluation. In addition, this study aims to find data sources that can be built for training artificial intelligence evaluation models, consistent with the importance of data acquisition and evaluation datasets mentioned in the literature (Liu & Kim, 2023). It shows that research focuses on data types and considers data quality and applicability, which is crucial for developing effective AI-aided design evaluation tools.

Overall, the design of this study fully considers the application prospects and challenges of artificial intelligence technology in product design, aiming to provide an important basis for developing AI-aided product design evaluation tools through systematic data collection and statistical analysis. This method not only finds the actual needs of designers but may also reveal the potential of AI technology in improving product innovation and design efficiency (Quan et al., 2023; Yüksel et al., 2023).

3.1.2.2 Natural Language Processing (NLP)

Natural language processing (NLP) has become valuable in various fields, including product design evaluation. NLP technology is used in the thesis in two stages of Study 3. First, NLP technology analyzes the relationship

between text data and scores in the dataset's construction stage. Using NLP methods; researchers can extract valuable insights from the feedback on the relationship between text and scores, providing important data for the next development of AI-assisted design evaluation tools. Secondly, in the algorithm implementation stage, Bert, an advanced NLP network, is a core part of the AI evaluation model. Researchers use the Bert algorithm to process and analyze the text description of product design.

NLP algorithms can analyze insights and extract unstructured text data (Kersloot et al., 2020) . This capability extends to areas such as mass customization in business, where text analysis from social media can influence product design decisions (Piriyakul et al., 2022) . In the context of product design, NLP can aid in tasks such as image retrieval through text processing analysis, proving the versatility of this technology (Piriyakul et al., 2022).

An important application of NLP in product design evaluation is the assessment of design documents to figure out factors such as risk levels, uncertainty of specifications, team temperament, and overall design process management (Wang & Dong, 2008) . By calculating and analyzing the semantic orientation of design text, NLP can provide valuable input for improving the quality and effectiveness of product concepts. In addition, NLP systems are crucial in capturing and standardizing unstructured clinical information, proving their ability to generate structured data from free text sources (Kreimeyer et al., 2017).

In healthcare, NLP has played a key role in improving the accuracy of automated processes, such as finding gout attacks through electronic health records (Yoshida, 2023). By combining NLP technology with health data, researchers have improved their recognition and understanding of medical conditions, showing the potential of NLP in refining healthcare product concepts. In addition, NLP methods are increasingly used to extract knowledge from unstructured health-related texts, emphasizing the importance of using NLP to gain insights from various sources (Gonzalez-Hernandez et al., 2017).

The practical application of NLP technology is crucial for supporting decision-making processes, as proved by knowledge-based operator-assisted support (Moghaddam, 2024). Using NLP capabilities, operators can receive help from enhanced tools that simplify their tasks and improve operational efficiency. In addition, NLP plays a crucial role in automating text analysis to evaluate the quality of student research work, showing its practicality in educational environments (Tarkhova, 2023). This automation simplifies the evaluation process and provides valuable feedback to students, helping to improve educational outcomes.

Haase et al. (2018) conducted a study on the effectiveness of advertising content in product promotion, writing down that the different dimensions of product design emphasized in advertisements can affect product evaluation. This highlights the role of text and visual elements in shaping consumer attitudes and preferences. In addition, evaluating students' multimodal texts in educational environments reveals the value of incorporating diverse representation patterns in learning materials (Ørevik, 2023). By recognizing the semiotic work involved in creating multimodal texts, educators can better

support students in expressing their ideas and understanding complex concepts.

In the context of design research, natural language processing (NLP) technology has become a valuable approach (Siddharth et al., 2022). NLP technology supports various aspects of design research, including analyzing textual data, extracting design-related information, and improving the overall design thinking process (Siddharth et al., 2022). Using NLP, researchers can gain deeper insights from textual data, enabling them to make informed decisions in the design evaluation process. Integrating NLP in design research marks a shift towards more advanced and technology-driven methods in the field.

In summary, the integration of NLP technology, language analysis, and user-centered design principles is crucial for evaluating and enhancing product concepts through text processing and design evaluation. These methods allow designers and educators to extract valuable insights, improve communication efficiency, and create attractive products that meet different user needs and preferences. Based on this, the study found the expert evaluation of multimodal product design projects' datasets that need to be collected. Natural language processing technology helps analyze a large amount of text data from product design projects, extracting valuable insights (Rath et al., 2022).

3.1.2.3 Computer Vision (CV)

Research has shown that the visual appeal of a product plays a crucial role in consumer decision-making processes (Creusen et al., 2010; Creusen &

Schoormans, 2004). Consumers often associate aesthetic product design with quality and value, making it a key factor in their purchasing choices. Therefore, accurately assessing and predicting the visual aesthetic quality of products can provide valuable insights for designers and businesses aimed at improving market competitiveness (Xia et al., 2016; Zheng, 2024).

Visual aesthetic quality assessment in product design is a key aspect that can significantly influence consumer preferences and market success. Therefore, the exploration of visual algorithms is the beginning of this study. Deep Convolutional Neural Networks (DCNN) have become powerful tools for analyzing and predicting aesthetic attributes in various fields. Researchers explored the application of DCNN in predicting aesthetic ratings and distributions based on product images (Jin et al., 2018; J. Wu et al., 2020). These models use the functionality of DCNN to extract complex features from images, enabling highly correct evaluation of visual aesthetics (Malu et al., 2017).

Using DCNN to evaluate and predict aesthetic attributes in product design may completely change the design process. By analyzing the visual features of a product, designers can obtain valuable feedback on the perceived aesthetic quality of their design, allowing for iterative improvement and optimization (Kuzovkin et al., 2017). This iterative process is consistent with the iterative nature of design thinking, in which continuous feedback and improvement are crucial for creating successful and visually appealing products.

The application of DCNN in aesthetic quality assessment is not only in product design but also in various fields such as photography, urban planning, and forestry. Researchers have developed models using DCNN to evaluate the aesthetic quality of various images (Chen et al., 2020; Ma et al., 2020; Mundher et al., 2022) . These applications highlight the versatility of DCNN in evaluating visual aesthetics in different fields and prove their potential in enhancing design related decision-making processes.

Furthermore, the correlation between the activation of DCNNs and human visual cortex activity emphasizes the biological relevance of deep learning models in understanding visual aesthetics. This consistency shows that DCNN can effectively capture and analyze visual features that resonate with human perceptual processes, thereby improving the reliability and effectiveness of aesthetic quality assessment.

In summary, integrating the Visual Neural Networks in visual aesthetic quality assessment has enormous potential to enhance the design process, improve consumer satisfaction, and drive market success through product design. The application of visual neural networks can systematically extract and interpret visual features, thereby promoting data-driven decision-making processes and accurately reflecting consumer behavior patterns and current market dynamics. The interdisciplinary application of Visual Neural Networks in aesthetic evaluation highlights their versatility and effectiveness in evaluating visual aesthetics in various fields, paving the way for innovative progress in design and decision-making processes.

3.1.2.4 Multimodal Neural Network

In product concept design evaluation, integrating neural networks, especially multimodal neural networks, has become a powerful tool to enhance the quality assessment process. Multimodal Graph Neural Architecture Search (MGNAS) is emphasized as a successful method for automatically designing the best architectures by combining multimodal representations, cross-modal information, and graph structures (Cai, 2024). This method allows for comprehensive analysis of various design aspects using multiple data patterns, thereby achieving more powerful and effective design evaluation.

Related studies have also proposed using deep learning techniques such as variational Onsager neural networks and Osprey optimization algorithms to evaluate product appearance aesthetics in industrial design (Zheng, 2024). By adopting advanced neural network models, designers can delve deeper into the aesthetic aspects of product design, ensuring that functional requirements are met, and that visual appeal and user experience are optimized. In addition, applying neural networks such as BP neural networks has played a significant role in product innovation design methods, promoting the construction of nonlinear mapping models between product attribute space and semantic space (Huang, 2022). By using the power of neural networks, designers can enhance a company's innovation capabilities, improve product competitiveness, and increase customer satisfaction through more refined and targeted design strategies.

The importance of multimodal deep learning in various fields, including activity and context recognition, has been studied, showing that multimodal neural network models are competitive compared to traditional methods (Radu et al., 2018). These findings emphasize the versatility and effectiveness of multimodal neural networks in handling various datasets and inference

tasks, highlighting their potential to advance product concept design evaluation.

Integrating morphological analysis theory and artificial neural network methods helps figure out the best product form and color combination during the design phase in sustainable product design (Hassan et al., 2012). By using neural networks, designers can make informed decisions about product aesthetics and functionality, ensuring the integration of sustainable design principles into the product development process. The integration of Kansei engineering with BP neural network algorithms in product pattern design systems facilitates enhanced comprehension of consumer perceptual preferences, thereby enabling enterprises to implement customized design solutions (Chen & Cheng, 2021). Integrating neural networks and design methods allows for a more systematic and targeted approach to product pattern development, effectively aligning design elements with user preferences.

Researchers use computer vision technology to process image data, combined with NLP algorithms to process and analyze the text description of product design and build a comprehensive multimodal AI evaluation model. This method of combining visual and text analysis can more comprehensively capture all aspects of product design, thereby improving the accuracy and effectiveness of evaluation.

3.1.2.5 Experiments

This study employed experimental design to evaluate the performance of the AI model. This includes human-machine comparative evaluation experiments,

allowing researchers to directly compare the performance of AI and human experts in design evaluation tasks (Rath et al., 2022). Through this method, the effectiveness and reliability of AI models are evaluated.

In the field of human-machine comparative evaluation experiments, a study delved into the complex and subtle differences in feeling between humans and machines, emphasizing that human errors are unlikely to be driven by the heuristic methods targeted in the research (Lepori & Firestone, 2020). This insight emphasizes the complexity of understanding the differences between human and machine cognition and highlights the necessity of meticulous evaluation methods. In addition, Testoni and Bernardi (2022) supplemented quantitative analysis by conducting human evaluation experiments on machine-to-machine dialogue, thereby contributing to this field, and demonstrating the importance of integrating human perspectives in evaluating machine interaction.

In terms of validating design prototypes, Ahufinger et al. (2019) proposed a user centered smartphone application for wireless electroencephalography in epilepsy patients, using expert group guidance, prototype design, and usability testing to confirm and validate the application's usage environment. The proposed method emphasizes the crucial role of integrating user input and expert evaluation in the design validation process. Research using a hybrid approach for design evaluation has shown that collaborative prototyping promotes problem-solving through iterative modeling during the initial product development phase (Bogers & Horst, 2013). This study emphasizes the collaborative aspect of design evaluation and proves how different perspectives can enhance the validation process. Carvajal et al. (2014) advocated using systems engineering methods for virtual/real analysis

and validation of automated systems, proving a shift towards early product validation of virtual prototypes. This transformation highlights the efficiency and effectiveness of virtual prototypes in validating design concepts. In the context of prototype validation, Nabilah (2024) focuses on creating a groundbreaking prototype for domain sentiment in project-based learning, emphasizing the importance of expert validation to ensure that the tool's content is aligned with its aims.

The synthesis of these references emphasizes the complexity of human-machine comparative evaluation experiments, design prototype validation, and the application of hybrid research methods in design evaluation. By integrating insights from those studies, multiple design evaluation methods, including human perspectives, experimental prototypes, and expert evaluation guidance, are crucial for ensuring the effectiveness and validity of design prototypes.

3.1.3 Mixed Methods: Questionnaire& Structured Interview

Combining questionnaire surveys and interviews has been widely adopted to evaluate user satisfaction effectively. Holmes et al. (1998) showed the effectiveness of this method in evaluating users' beliefs about psychiatric day hospitals, obtaining data on overall satisfaction scores, specific aspects of service satisfaction, finding issues, and providing improvement recommendations (Holmes et al., 1998) . Similarly, Lobo et al. (2013) conducted interviews with over 6000 primary healthcare service users to assess satisfaction levels and constructed a global satisfaction index using a partial least squares path model. These studies emphasize the importance of

using questionnaire surveys and interviews to collect comprehensive data on user satisfaction.

In the field of technology, Kanchymalay et al. (2013) focused on ERP customization for manufacturing companies, using in-depth interviews and surveys to measure user satisfaction in daily operations. This study emphasizes the importance of understanding user perspectives through direct interaction in improving satisfaction levels. In addition, Miranda (2024) also surveyed users' satisfaction with the regional health comprehensive clinic in Bahia and Brazil. A combination of questionnaire surveys, semi-structured interviews, and observations was used as a method to collect quantitative and qualitative data for overall evaluation. This comprehensive approach allows for a comprehensive evaluation of user satisfaction and service quality.

In addition, Boiani et al. (2019) examined the satisfaction of elderly users with mobile assistive devices through interviews and Quebec user satisfaction assessment questionnaires, emphasizing the importance of direct user feedback in evaluating device satisfaction. Okafor (2024) emphasized the impact of a user-centered design approach on improving user satisfaction and experience, saying that focusing on usability can significantly enhance user feeling. Similarly, Wu (2024) explored the factors affecting user satisfaction with online agricultural products in Chinese cities during the pandemic, emphasizing optimism and product quality in improving user satisfaction. These studies emphasize the diverse applications of user satisfaction assessment in different industries and backgrounds.

In summary, the study of user satisfaction through questionnaire surveys, structured interviews, and user experience evaluations reveals the multidimensionality of user feedback. Especially in the field of artificial intelligence usability evaluation, by combining quantitative and qualitative methods, researchers can comprehensively understand user feelings, find areas for improvement, and improve user satisfaction.

Questionnaire surveys collected a large amount of standardized data (Tubadji et al., 2021), which is crucial for understanding the everyday needs and attitudes of the designer community. In design research, questionnaire survey methods play a crucial role. Questionnaire surveys are regarded as the primary method for collecting data in various fields, including academic research, public institutions, and business organizations research and data collection work (Sheth, 2010). It involves meticulous planning, construction, and writing of survey materials to elicit expected responses from participants (Sheth, 2010; Taherdoost, 2022). In addition, questionnaire surveys are quantitative and qualitative research methods to collect participant insights (Hendriana et al., 2023; Ohueri et al., 2019).

In the research of survey methods, questionnaire design is a key part that affects the quality and reliability of the collected data (Rattray & Jones, 2007). Researchers emphasize the importance of using pre-planned methods to determine the questionnaire's reliability and validity to ensure the collected data's robustness (Rattray & Jones, 2007). In addition, the questionnaire design should aim to reduce potential biases and errors that may affect the research results (Antoun et al., 2017). By following best practices in questionnaire design and development, researchers can improve the rigor and credibility of their research results.

During the questionnaire survey process for product design evaluation, researchers must consider a range of factors to ensure the effectiveness of the data collection process. Choosing survey design, constructing survey tools, and evaluating reliability and effectiveness are key stages of survey-based research (Kitchenham & Pfleeger, 2008). Researchers must also pay attention to the management of survey tools and next analysis of collected data to obtain meaningful insights (Kitchenham & Pfleeger, 2008). By following the systematic questionnaire survey design and data analysis methods, researchers can obtain valuable information to provide insights for product design decisions.

Questionnaire surveys provide a structured method for collecting feedback and insights when analyzing data obtained from expert product design evaluations. Using questionnaire surveys enables researchers to systematically collect data on experts' opinions, preferences, and criticisms of product design (Yang, 1989). The data obtained from the questionnaire can be comprehensively statistically analyzed, promoting evidence-based evaluation of product design quality and effectiveness (Cai, 2024).

In summary, the questionnaire survey method plays a significant role in design research, especially in evaluating product design through expert data analysis. By using carefully designed questionnaires, researchers can improve the quality of data collection and analysis in design studies. Through meticulous planning, structured survey material design, and rigorous data analysis, researchers can gain valuable insights, provide information for product design decisions, and contribute to progress in the field.

Methodologically, using the questionnaire survey method and SPSS for statistical analysis is a scientific and effective research method. This method can collect the actual needs and perspectives of designers, which is in line with the emphasis in the literature on combining artificial intelligence technology with people's actual needs (Liu & Kim, 2023). A questionnaire survey can help understand how designers view the application of AI technology in fresh design stages, which is of great significance for defining product design stages that require AI-assisted evaluation.

In addition, this study aims to find data sources that can be built for training artificial intelligence evaluation models, consistent with the importance of data acquisition and evaluation datasets mentioned in the literature (Liu & Kim, 2023). It shows that research focuses on data types and considers data quality and applicability, which is crucial for developing effective AI-assisted design evaluation tools.

Overall, the design of this study fully considers the application prospects and challenges of artificial intelligence technology in product design, aiming to provide an important basis for developing AI-assisted product design evaluation tools through systematic data collection and statistical analysis. This method not only finds the actual needs of designers but may also reveal the potential of AI technology in improving product innovation and design efficiency (Quan et al., 2023; Yüksel et al., 2023).

3.1.4 Reasons for the Mixed Methods

There are three main reasons for choosing a mixed methods approach. First, it provides a deeper understanding of complex research questions from

multiple perspectives. Second, triangulation can improve the reliability and validity of research findings. Finally, qualitative methods provide in-depth insights, while generalizing experimental results through quantitative methods can enhance understanding (Gezici & Tarhan, 2022; Rosen et al., 2023). This combination of methods not only allows for a comprehensive analysis of the problem, but also allows for the exploration of nuances of meaning and process during the data collection and analysis phases (Charmaz, 2015). Mixed methods research is increasingly being used in applied health research, with many researchers advocating the use of both qualitative and quantitative methods (Rosen et al., 2023). This approach is particularly well suited to the study of emerging phenomena because it can effectively deal with the lack of existing theory and uncover useful and important patterns and concepts through continuous comparison of data (Glaser et al., 2005).

This thesis adopts a mixed research method to understand complex research problems from multiple perspectives. Qualitative methods uncover the real needs of designers and users in the AI era, gaining in-depth insights into the development of AI products. Quantitative methods enable the validation of AI product design evaluation models, thereby promoting the theoretical framework for improving design evaluation methods using AI technology. The reliability and effectiveness of research results are enhanced through triangulation, using multiple data sources and analysis methods to enhance the credibility of research findings.

This study adapts to the interdisciplinary nature of the research question. AI-assisted product design evaluation involves multiple fields, such as technology, user experience, and design theory, requiring a multifaceted

research method. Conducting research through mixed methods provides a comprehensive research perspective, covering everything from requirement analysis to model development and practical application evaluation. Eventually, those methods enhance the practicality and generalization of research results. Combining theoretical exploration and practical verification makes the research results more likely to be applied in practical design environments. Through this systematic and multimethod research strategy, this study aims to comprehensively explore the application of AI in product design evaluation, providing a solid theoretical foundation and practical guidance for this field. Through rigorous data collection and analysis processes, insightful qualitative research results are expected to develop (Charmaz, 2015; Rosen et al., 2023). This method not only develops innovative AI-driven product design evaluation models but also gains a deeper understanding of these technologies' application effects and potential impacts in practical design environments (Vianello et al., 2023).

3.2 The Methodology Framework

Through a literature review, the current research gaps were found and defined as RQs 1-4. At the same time, it suggested potential methods or techniques for AI-assisted product design evaluation. This study employed a mixed approach that combines qualitative and quantitative techniques to comprehensively examine the multifaceted applications of artificial intelligence in the design evaluation process. This method, based on the multidimensional nature of the research problem, can provide deeper and more comprehensive insights (Charmaz, 2015; Rosen et al., 2023) . The research design is divided into four main stages, each corresponding to a research question (RQ). The research is structured into four main phases, each corresponding to a research question (RQ), as shown in Figure 3.1.

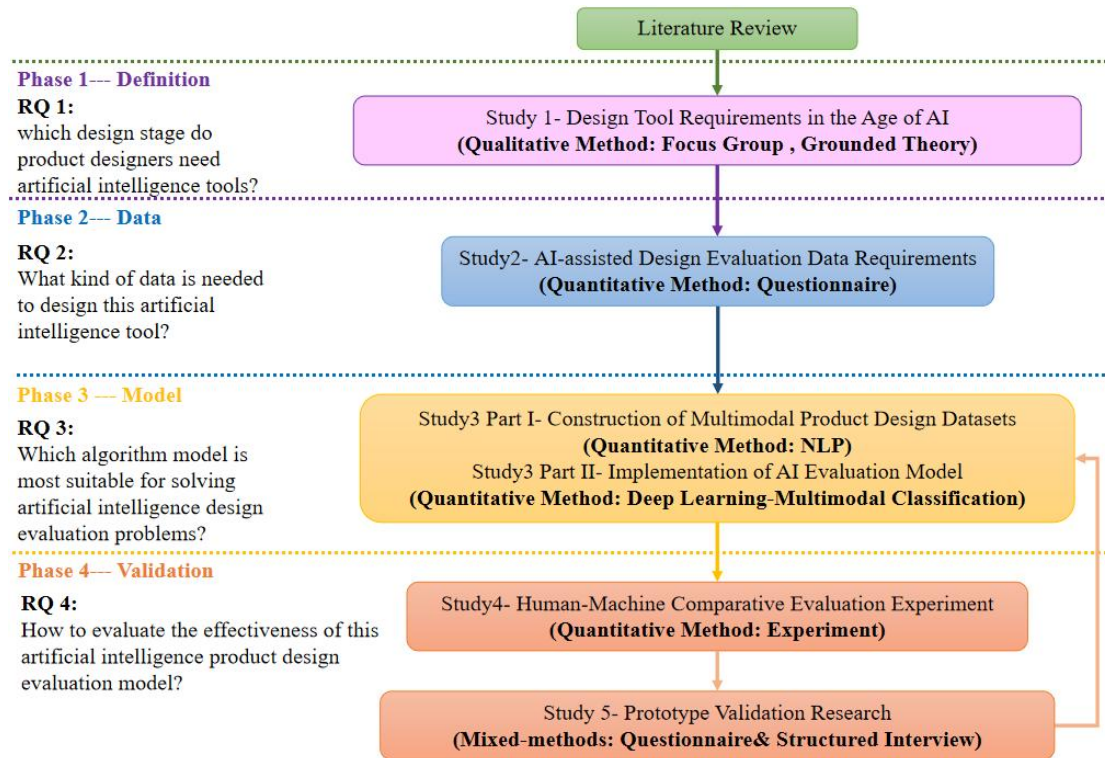


Figure 3.1 The Framework of Methodology

3.2.1 Definition

In this stage, qualitative methods are used to explore the tool needs of product designers in the AI era through focus group interviews and grounded theory analysis. This stage aims to understand the designers' actual needs better and lay the foundation for the following research (Charmaz, 2015; Rosen et al., 2023). This study invited design students and experts from various backgrounds, including representatives from academia and industry (Lopez & Garza, 2023). The focus group discussion structured into multiple blocks. It covers various stages of AI application in the design process. This method allows us to understand the designers' actual needs better, laying the foundation for the following research. In this study, focus group interviews were strategically employed as a primary method for collecting qualitative data (Charmaz, 2015). This approach eased in-depth discussions among

diverse participants, including designers, technologists, and end-users, who shared their experiences and perspectives on integrating artificial intelligence (AI) in design processes.

The data gathered from these focus groups was thoroughly analyzed using grounded theory (Glaser et al., 2005). This well-established qualitative research method aims to construct theories through systematically collecting and analyzing data. Grounded theory is particularly valued in fields where research is aimed at uncovering how individuals interpret their experiences, making it an excellent choice for studying the nuanced impacts of AI on design practices. NVivo, a sophisticated data analysis software, was used to code and analyze the interview data. NVivo supports researchers in managing and analyzing large volumes of text-based data, enhancing the efficiency and depth of qualitative research (Feuston & Brubaker, 2021). The coding was conducted at three distinct levels (Glaser et al., 2005) . Open Coding: The data was meticulously examined to show initial codes at this foundational stage. This involves labeling data segments with a short name that accurately describes the content. Open coding is crucial for understanding the breadth and depth of content discussed during the interviews. Axial Coding: Building on the open codes, axial coding was used to begin linking these codes together into categories based on their relationships. This middle step focuses on the connections between the abstracted data, helping to form a more coherent and integrated understanding of the central phenomena being studied. Selective Coding: The final stage of the selective coding process involved refining and integrating the categories to form a comprehensive narrative. This stage focuses on finding the consistent core themes throughout the data, which serve as the central concepts for the emerging theory.

This structured approach to data analysis is particularly effective for exploring the emerging applications of AI in the design field. By employing grounded theory and NVivo, the study not only ensures a rigorous analysis of qualitative data but also eases the extraction of meaningful insights deeply embedded in the real-world experiences of the participants (Feuston & Brubaker, 2021). Multilevel coding allows researchers to distill complex data into understandable and actionable concepts and theories (Glaser et al., 2005). These insights are crucial for understanding both AI's current landscape and future potential in enhancing design processes. They offer valuable guidance for developers creating AI tools tailored to the needs of designers and for designers looking to use AI to enhance their creative workflows. The findings from this research can help bridge the gap between technological capabilities and user needs, leading to more innovative and effective design solutions.

3.2.2 Data

At this phase, research methods shift towards quantitative methods and transform qualitative insights into quantifiable data requirements (Neto et al., 2018). First, a questionnaire survey was used to decide the data types and sources needed for designing AI-assisted tools. The questionnaire design is based on the findings of the first stage and reviewed by researchers to ensure its validity (Neto et al., 2018). After collecting the datasets, natural language processing (NLP) techniques are used to analyze the keywords and indicators of expert evaluation data, which helps us understand the structure and characteristics of design data (Rath et al., 2022), and analyze the data types and sources required for designing AI-assisted tools. This study phase used a structured approach to gather quantitative data for evaluating product design. A questionnaire survey method was implemented to capture various data

points relevant to product design evaluation. These included metrics on usability, functionality, and user satisfaction. The collected data provided a quantitative foundation for further analysis and insight generation.

SPSS software was employed to manage and analyze these substantial datasets (Shao et al., 2021). Renowned for its robust statistical analysis capabilities, SPSS eased a thorough data exploration. It enabled the execution of complex statistical tests and offered reliable insights into patterns and correlations within the data. The software's advanced analytical tools, such as regression analysis, variance analysis, and significance testing, were instrumental in quantifying the relationships between design elements and their impact on user feeling and usability (Shao et al., 2021). It integrated the quantitative data analysis with SPSS.

During Study 2, a comprehensive collection of multimodal product design datasets was amassed, encompassing a diverse range of image and text data. These datasets were specifically curated to reflect the multifaceted nature of product design, capturing everything from visual images and functional specifications to textual descriptions and experts' scoring. Survey-derived textual data was computationally processed using natural language processing (NLP) methodologies, serving as a valuable adjunct to the numerically-based examination. The BERT (Bidirectional Encoder Representations from Transformers) algorithm, a state-of-the-art NLP model, was used for its ability to understand the nuances of human language (Dosovitskiy, 2020). By employing BERT, the study used its capabilities for word frequency analysis and clustering analysis (Dosovitskiy, 2020).

The word frequency analysis involved parsing copious amounts of text to find and quantify the most often used terms and phrases. This process helped highlight common themes and concepts discussed by respondents, offering a direct insight into prevalent trends and user priorities in product design. Clustering analysis, on the other hand, grouped similar text responses together, enabling the identification of underlying patterns and categories within the textual data. This analysis was crucial for segmenting the data into meaningful clusters that reflected distinct aspects of user feedback, such as specific design features or common issues met. The qualitative insights derived from NLP techniques allowed the study to understand the numerical and textual data comprehensively. This dual approach allowed for a more nuanced analysis, where quantitative findings could be enriched and contextualized with qualitative data, offering a holistic view of the structure and features of design data. The insights gained from this phase were instrumental in informing the next phases of the study, particularly in refining the AI models used for product design evaluation. The combination of SPSS and BERT provided a robust method for dissecting and understanding the complex interplay of factors contributing to successful product design.

3.2.3 Modeling

This phase involves building multimodal product design datasets and implementing an AI evaluation model. Deep learning and multimodal classification methods were adopted to address this research's core technical challenges directly. The construction of the datasets considers multiple modalities, including image and text data, to comprehensively capture the complexity of product design. Build multimodal product design datasets and design and implement an AI evaluation model. The model adopts deep

learning and multimodal classification methods to directly address the core technical challenges of the research (Neto et al., 2018). The model training process includes hyper-parameter tuning and model selection to ensure the best performance (Rath et al., 2022). To harness the full potential of datasets, they were subjected to training and evaluation using a variety of sophisticated deep-learning algorithms. Among the visual algorithms employed were Vgg16 (Pal et al., 2023), ResNet18 (He et al., 2016), DenseNet121 (Bulbul et al., 2022), CustomNet (Madni, 2024), and Vision Transformer (ViT) (Dosovitskiy et al., 2020). Each of these algorithms offers unique strengths in processing complex visual information, making them well-suited for analyzing the visual components of the product designs.

Additionally, the language understanding capabilities of the BERT model (Devlin, 2018) were used to interpret and analyze the textual data going with the product designs. BERT's ability to process natural language and extract meaningful patterns is crucial for understanding text's descriptive and functional aspects. Multimodal combination models were also evaluated to fully exploit the synergies between visual and textual data. These included integrations of Vgg16 & BERT and ResNet18 & BERT (Shao et al., 2021). These combination models aim to merge the insights from visual and textual analysis, offering a holistic view of the product designs that mirror how humans perceive and evaluate products.

3.2.4 Validation

The final phase consists of two main parts: human-machine comparative evaluation experiments and prototype validation research. It comprehensively evaluates the effectiveness of AI models using a

combination of quantitative experiments and mixed methods (Christou, 2023; Neto et al., 2018). The experts and AI models simultaneously rated samples in the human-machine comparison experiment. Then, the rating results are compared to evaluate the accuracy and reliability of the model (Lopez & Garza, 2023) . Prototype validation research collects user feedback on the model's prediction results through questionnaires and structured interviews. It aims to understand users' views on the model's satisfaction, accuracy, and usability (Tubadji et al., 2021) . Study 4 implements an innovative online experimental methodology, inviting designers from various fields to digitally submit their product design projects. This approach allows for a broad range of design styles and innovation levels to be considered, providing a diverse dataset for evaluation.

These projects are evaluated by a panel of experienced experts who conduct their assessments online. This method ensures that the evaluation process is efficient and scalable, allowing for a larger number of evaluations in a shorter period. Each project is assessed based on a set of predefined criteria. Simultaneously, sophisticated AI models evaluate the same projects. These models are designed to analyze and rate the projects using advanced algorithms that mimic human evaluation processes. The AI models have capabilities such as image recognition, text analysis, and pattern detection to provide comprehensive assessments.

The core of this study is the comparative analysis between the ratings provided by human experts and those generated by AI models. This comparison is eased with algorithmic statistics, which provide a quantitative measure of agreement or discrepancy between human and machine

evaluations. This analysis helps find consistent patterns or significant evaluation criteria or outcome differences.

This approach aims to evaluate the accuracy and reliability of the AI models rigorously. The study determines how well AI can replicate human cognitive processes in product design evaluation by comparing AI-generated results with expert human judgments. Furthermore, this experiment explores the potential for AI to aid or even augment human evaluators in the future, particularly in terms of efficiency and the ability to manage large datasets. Additionally, the outcomes of this study are expected to contribute to the ongoing discourse on integrating AI into creative processes. Insights gained from the discrepancies and alignments between human and AI evaluations could improve AI algorithms, making them more sensitive and aligned with human aesthetic and functional judgments.

The quantitative part of this phase involved questionnaire surveys, which were systematically designed to gather measurable data from users. These surveys had structured questions that quantified participants' beliefs of the AI model's performance and usability. The responses to these questionnaires were analyzed using SPSS (Statistical Package for the Social Sciences), a powerful software tool renowned for performing complex statistical analysis (Shao et al., 2021) . SPSS enabled the research team to conduct a variety of statistical tests and reliability testing, which provided insights into the correlations, trends, and consistencies within the quantitative data. This statistical analysis was essential for objectively measuring the AI model's accuracy and user satisfaction, providing clear metrics to quantify the model's effectiveness. In this study phase, a mixed-method approach was strategically employed to comprehensively evaluate the AI evaluation model, specifically

focusing on user satisfaction, accuracy, and usability. This approach combined the strengths of both quantitative and qualitative research methods to ensure a thorough analysis of diverse data types.

Complementing the quantitative analysis, the qualitative data collected from structured interviews offered deeper insights into the users' experiences and beliefs. These interviews allowed participants to discuss their views and experiences in a more open-ended format, providing rich, narrative data that could reveal nuances not captured by the quantitative survey. The qualitative data was analyzed using NVivo, a sophisticated qualitative data analysis software that helps the organization, coding, and thematic analysis of textual data (Dlugatch et al., 2023). Through NVivo, the research team performed detailed content analysis, finding recurring themes and patterns related to the AI evaluation model's usability and user satisfaction aspects. This qualitative analysis helped contextualize the statistical findings, adding depth to understanding how users interact with and perceive the AI model. Integrating findings from both SPSS and NVivo provided a robust evaluation of the AI evaluation model. The study achieved a balanced view of the model's performance by combining the precise, quantitative data from the surveys with the nuanced, qualitative insights from the interviews. This mixed-method approach confirmed the quantitative findings with qualitative evidence and highlighted any discrepancies that could write down areas for further improvement.

This comprehensive evaluation was crucial for developing a well-rounded understanding of the AI model's effectiveness in real-world settings. It ensured that the model met statistical performance standards and met user expectations and needs in practical applications. This multimethod and

multi-stage research design can comprehensively explore AI's application in product design evaluation. This study covers the entire process of AI-assisted design evaluation, from user needs analysis to model development and practical application evaluation. This method offers in-depth theoretical insights and ensures the practical relevance of research results.

Those studies provide the methodologies used for algorithmic comparison and the statistical techniques that enhance the understanding of AI capabilities in mimicking human evaluation (Feuston & Brubaker, 2021; Lange, 2024) . There is a logical connection between these four phases. The discovery during the definition phase guided the requirement analysis during the data phase. The results of the data phase provide the necessary inputs for the model phase. The validation phase involves comprehensive testing and evaluation of the entire AI-assisted evaluation system. This coherent research process ensures that research questions are systematically and comprehensively answered.

3.3 Summary of Chapter Three

This methodology is based on theoretical foundations and empirical efficiency, aiming to systematically explore, implement, and confirm artificial intelligence-driven conceptual product evaluation models. This model excels at handling and analyzing different dataset features in multimodal conceptual product design. The core of this method is an intelligent prototype that excels in data collection, preprocessing, and visual evaluation of data. This prototype aims to adapt to various aspects of product design, including visual, textual, and functional patterns. This thesis used multimodal artificial intelligence models to address the challenges posed by such rich datasets,

supplemented by comparative analysis with human evaluation metrics. This examines the performance and potential of artificial intelligence in this new context and plays a certain role in explaining the methodological progress of multimodal data in design research.

Studies 1- Study 5 form a methodological framework for using artificial intelligence technology to improve product design evaluation methods, emphasizing the potential of AI in improving evaluation efficiency and accuracy.

Study 1 directly answers RQ1 through focus group interviews and grounded theory methods. It delves into the specific needs of designers in the AI era, providing a foundation for future research. This method can capture the designers' thoughts and expectations, which helps develop practical AI-assisted tools.

Study 2 uses questionnaire surveys and natural language processing techniques to answer RQ2. It systematically analyzes the multimodal data types and sources needed for design evaluation. This stage of research not only found key data types but also explored best practices for data collection and integration.

Study 3 involves training the multimodal product design datasets, designing, and implementing an AI evaluation model, and answering RQ3. This phase involves advanced machine learning and deep learning techniques such as multimodal fusion and transfer learning. By comparing the performance of

different algorithms, this stage provides a scientific basis for selecting the most suitable AI model for product design evaluation.

Study 4 and Study 5 comprehensively answer RQ4 through human-machine comparative evaluation experiments and prototype verification studies. These two research phases evaluated the technical performance of AI models and delved into their practical application effects in design environments. By collecting and analyzing user feedback, these studies offer valuable insights for further optimization and practical deployment of AI models.

The purpose of providing a detailed introduction to the methodology is not only to provide a clear roadmap for this thesis but also to offer an applicable template for the intersection of artificial intelligence and product design. This research design ensures that each research question is systematically and multimethodologically explored. By combining qualitative and quantitative methods, this study can develop innovative AI-driven product design evaluation models and provide a solid theoretical foundation and practical guidance for applying AI in the design field. This comprehensive approach helps bridge the gap between academic research and industrial applications, promoting the effective application of AI technology in product design evaluation. Although this paper still has limitations, it adopts a rigorous research method to deal with emerging phenomena and provides a methodological framework for designing and developing artificial intelligence models. Recognizing these limitations also provides directions for improvement and expansion for future research. The next studies (Study 1-Study 5) will be performed under this methodological framework.

This chapter elaborates on the study's methodological framework, data collection, and analysis methods, laying the methodological foundation for the following five interrelated studies (Study 1- Study 5). This thesis ensures the comprehensiveness and reliability of the study by designing a multimethod, multi-stage research framework. This methodological design directly guides the implementation of Study 1- Study 5 and provides a specific method for exploring designers' needs for AI-assisted design tools. Although this chapter proposes a comprehensive methodological framework, its effectiveness still needs to be verified through the following research.

Chapter 4. Study 1- Design Tool Requirements in the Age of AI

This chapter introduces Study 1, which aims to define designers' needs for AI-assisted design tools. Based on the guidance of the methodological framework in the earlier chapter, this study determined to use focus groups and grounded theory methods to collect and analyze qualitative data from designers from academia and industry. NVivo software was used to process the data using a three-level coding method based on grounded theory, and the requirements for AI-assisted design tools were found. Among them, it was found that AI design evaluation is an important requirement point. This study offers key insights into understanding designers' needs and expectations, laying the foundation for data collection and model development in the next studies. The findings in this chapter will directly influence the design of the questionnaire and the determination of data requirements in Study 2.

4.1 Introduction

Product designers primarily utilize artificial intelligence tools during the design and engineering phases of the product lifecycle. These tools are crucial for enhancing creativity and efficiency in the design process (Figoli et al., 2022). Using artificial intelligence, designers can access various applications such as scheduling, document management, project progress tracking, error detection, and correction to prevent key issues (Maksoud& Ahmed, 2024; Matter, 2024). In addition, AI image generation tools such as DALL • E2 and

Imagen allow designers to train text, graphics, and models, changing traditional design methods (Hu, 2024).

Artificial intelligence is increasingly integrated into architectural design processes, providing capabilities to simulate human intelligence in tasks such as analyzing designs, identifying errors, and suggesting alternative solutions (Maksoud& Ahmed, 2024) . Furthermore, AI technologies are crucial in optimizing plant layout designs, allowing designers to streamline workflows and quickly adapt to new product requirements (Shiralkar, 2017) . In industrial design, AI algorithms contribute to cognitive thinking storage and artistic image optimization, enhancing the overall product design process (Liu, 2015). In environmental design, artificial intelligence tools are evolving rapidly to enhance design methods, expressions, and content, leading to significant transformations in design language and creativity (Hou & Xu, 2021) . Moreover, AI's role in precision marketing within the new retail era focuses on improving user experiences and aligning marketing strategies with consumer demands through intelligent service scenarios (Lin & Zhang, 2019). This underscores the diverse applications of AI not only in product design but also in marketing strategies to meet evolving consumer needs.

The application of artificial intelligence in product design extends to digital multimedia technology, where interactive AI-based design systems facilitate the conversion of raw data into actionable design information, enhancing collaboration between designers and intelligent systems (Liu, 2022) . Additionally, AI-powered software autonomously approaches architectural design processes comprehensively, from upstream conceptualization to downstream implementation, highlighting AI's autonomy and efficiency in the design domain (Satwiko & Michelle, 2022). Artificial intelligence plays a

pivotal role in evaluating design sketches based on designers' cognition, selecting feasible solutions, and reconstructing product image cognitive systems, demonstrating the integration of AI in enhancing the design evaluation process (Yang et al., 2021). Furthermore, the use of generative artificial intelligence tools in education, as highlighted in the instructional design matrix approach, underscores the potential of AI in transforming educational practices through innovative design methodologies (Ruiz-Rojas, 2023).

Overall, artificial intelligence tools are essential in the design phase of product development, providing designers with a wide range of capabilities to enhance creativity and efficiency, optimize workflows, detect errors, and propose alternative design solutions. The integration of artificial intelligence in various design fields, such as architecture, industrial design, and environmental design, demonstrates the revolutionary impact of artificial intelligence on traditional design processes, promoting more innovative and efficient design practices. Therefore, exploring designers' perspectives and needs on artificial intelligence has significant research value. This is also the significance and main research objective of conducting Study 1.

4.2 Methods and Results

By gaining a deeper understanding of the needs of designers at various stages of design, subsequent research can develop more targeted and effective AI-assisted tools. The development of this study can promote human-machine collaboration and improve design efficiency and innovation.

4.2.1 Focus Group Discussion

Study 1 used the focus group method to interview 87 product designers, including 3 groups of 18 people from industry and 14 groups of 69 students from product design majors. This study discovered the design insights of these groups through focus group interviews. From a qualitative perspective, provide a framework of user needs and product functionality for Study 2, the data phase.

There are three stages for Focus group interviews. The first stage is the mining of user needs. The second stage is user experience interviews. The third stage is the perception and attitude of users towards innovative technologies.

In the first stage, participants use 5-10 minutes to fill in the following questions and then start the discussion and exchange. Each person will speak for about 5 minutes per question. The duration of the first stage is about 1 hour. Participants rest for 5 minutes. In the second stage, participants fill in the following questions during the break based on the content of the previous discussion. Based on the previous discussions and exchanges, each person will speak for about 5 minutes per question. The second stage lasts about 1 hour. Participants rest for 5 minutes. In the third stage, based on previous discussions, each participant will speak for about 5 minutes per question. The third stage lasts about 1 hour. Each focus group interview totaled about 2-3 hours. The details are in Table 4.1.

Table 4.1 Focus Group Outline

Stage	Questions	Intention
1	Q1: What are the processes or stages in the general product design project process, and what design tools?	<i>User Demand Mining</i>

	Q2: At various stages, are there tasks that are repetitive, cumbersome, and hope to be assisted by more people?	
2	Q1: Do you understand artificial intelligence? Based on your understanding, do you think artificial intelligence can solve the first stage of the problem?	<i>User Experience</i>
	Q2: Have you used smart design tools? What kind of smart tools are?	
3	Q1: What do you think of artificial intelligence-assisted design? Willing to actively embrace innovative technologies, or are there worries?	<i>Perception of AI Technology</i>
	Q2: Do you think that in the future designers can work with artificial intelligence to jointly develop new products? If so, how do you think it works together?	

This study used stratified sampling to select experimental participants to ensure the representativeness and diversity of the sample. The participants are mainly divided into two categories: the senior designer group (18 people) and the junior designer group (69 people). This layered design is based on the following considerations.

The reasons for choosing a senior designer are as follows. Firstly, they have rich practical experience and are familiar with the entire process of product development. Secondly, they have an interdisciplinary perspective and can provide multidimensional opinions. The most important thing is that they have market insight and can evaluate the commercial value of design solutions.

The reasons for choosing a junior designer are as follows. Firstly, they represent the main user group of design tools. Secondly, they have a high acceptance of new technologies and are more willing to try AI assisted tools.

And they can provide tool optimization suggestions from the user's perspective. Finally, their design thinking is more active and can test the innovation support ability of AI tools. This multi-level and multi-dimensional participant selection strategy ensures the scientific and representative nature of experimental data, providing a guarantee for the reliability of research results.

There are two participant categories to collect the opinions of target users more objectively. Among them, three groups are senior designers with 2-3 years of experience in a professional design company (Figure 4.1). The other group members are junior designers and third-year undergraduates in product design.



Figure 4.1 Focus Group- Senior Designers

A total of three focus group interviews were conducted and interviewed three professional design companies. Each of these three companies has its areas of expertise. One is focused on industrial design. The other is to provide an experience design. One also provides integrated design services (including product design and brand design). The details are in Table 4.2.

Table 4.2 Participants- Senior Designers

Company	Gender	Working years	Position
MOMA(Shenzhen)	M	5	Co-Founder
	M	4	Design Manager
	M	2	Industrial Designer
	M	2	Industrial Designer
	F	2	Industrial Designer
MOMOUX	M	3	Strategy Designer
	M	2	UI Designer
	M	2	UI Designer
	F	2	UX Designer
	F	2	Design Manager
	F	4	UI Designer
XIVO Design	M	5	Design Director
	M	5	Design Manager
	M	5	Industrial Designer
	M	3	Industrial Designer
	M	2	Industrial Designer
	M	4	Brand Designer

The main members of the junior designer group are students in the third year of product design. They come from a university in Guangdong Province, China. A total of 69 participants conducted fourteen focus group interviews. It is shown in Figure 4.2.



Figure 4.2 Focus Group- Product Design Students

4.2.2 Grounded Theory Analysis

The grounded theory three-stage coding method was used in this study. It was employed as the analytical approach for qualitative data coding, and NVivo was used for data analysis. Through Open Coding, Axial Coding, and Selective Coding, standard features or related meanings of cases or events are extracted to group them (Charmaz, 2015).

This study invited three researchers to code the focus group data to ensure the objectivity of the Coding. These three researchers all have research experience. They come from product design, user experience, and information technology. This study conducted reliability testing on three coders. Three researchers randomly selected 10% of the focus group data for independent coding and evaluated initial consistency using Fleiss' Kappa coefficient. Inter-coder reliability was assessed using Fleiss' Kappa to measure agreement among the three researchers. Coding results were exported from NVivo and analyzed in ReCal, and a Kappa value of 0.72, indicating substantial agreement. Given the multidisciplinary background of the coding personnel,

0.72 is an acceptable threshold, indicating strong consistency and high research reliability of the independent coding results of the three researchers on the data (Freelon, 2013). Researchers reached a consensus by discussing and revising the coding manual and unifying standards for inconsistent coding entries. During the formal coding process, cross sampling verification is repeated for every 20% of the data volume completed to ensure reliability stability. Moreover, researchers follow a unified coding process and guidelines, as shown in Figure 4.3.

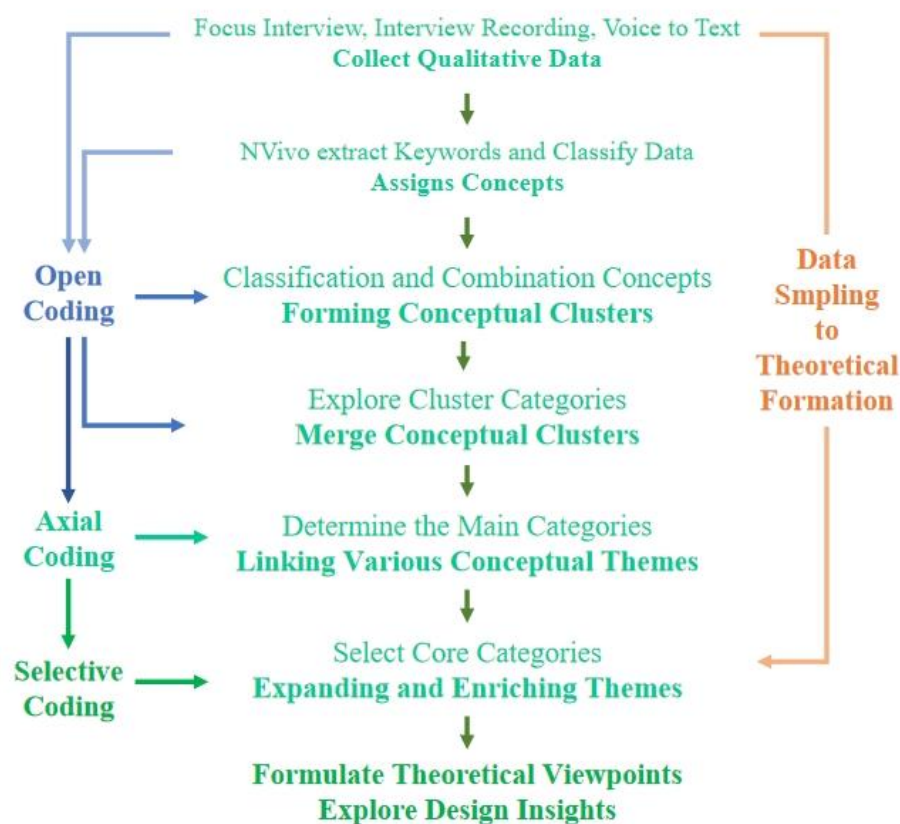


Figure 4.3 The Coding Flowchart

4.2.2.1 Open Coding

In the first-level Coding of interview data, the concept category is searched from the collected data. The keywords are extracted and labeled for

conceptualization. The following three methods are used to name the concept in the open coding stage.

(1) Self-creation: The coders conduct continuous comparative analysis of data, develop conceptual labels to capture the fundamental meanings or representations triggered by the studied phenomena.

(2) Use existing names: This naming is very rigorous because the concept itself in the authoritative literature contains wealthy analytical meaning, and they are almost complete concepts, so they are very rigorous, but the disadvantage is that they need flexibility. It may be different from the concept researchers want to express.

(3) Code in vivo: The coders systematically derive coded labels directly from the word for word expressions used by research participants in interviews.

The details are in Table 4.3.

Table 4.3 Open Codes

Clusters	Concepts	Original sentence examples (e.g.)
Standard Formulation	Template, Parametric Design, Font Size	1. It will provide you with a reference such as a center line or a three-point line. 2. Parameterized design. 3. The fonts are either Microsoft Yahei or regular Song typeface.
Design Tools	Microsoft PowerPoint, Plugin, Initial Template, Non Electronic Tools, Sticky Notes, Collaborative Office Software, Graphite Drawing Desk, Mobile Memo, Dual System	1. Mobile memo apps have collaborative office software (Mobile memo, collaborative office software) 2. Sometimes the layout of PPT also has this issue. (Microsoft PowerPoint) 3. It does not have a simple initial template, and then it has to be manually dragged over one by

		one (Initial Template)
Design Process	POV, Case Study, Background Research, Market Research, Emerging Technologies, Data Classification, Information Collection, User Research, Sketches, Planning, Product Experience, Initial Version, Iterative Testing, Interviewee Fit, Interview, Interview Content Repetition, Interview Questionnaire, Online Interview, CAD Modification, Storyboard, Modeling, Design Review, Design Insights, Design Element Extraction, Problem Setting, Physical Model, Visual Presentation, Pain Point Analysis, Brainstorming, Project Management, Information Visualization, Information Integration, Prototype Testing	1. I will write a process that includes background research, interview questionnaires, pain point identification (POV), draft modeling and rendering, and iterative testing. (Pain Point Analysis, Background Research, Iterative Testing, Interview Questionnaire) 2. We still need to collect case studies, and sometimes we may look for common things. (Case Studies) 3. Because there may be technical research in the early stage, for example, if you need to do something that can only be described as relatively technical or the latest materials, or about materials or new technologies, sometimes you may accidentally come across or not have a systematic understanding. (Emerging Technologies) 4. The mechanism for evaluating artificial intelligence may also be faster. (Design Review) 5. So decision-making is about choosing a topic, and this requirement may provide you with a problem to solve. But it is up to you to decide how to solve it. For example, designing the functionality of a product within the broader context of design. (Setting Questions)
Data	Big Data, Quantification, Data and Opinions, Search Engine Priority, Following Preset or Research Data	1. The human brain is not capable of thinking, but machines can process big data. (Big Data)

		2. Is it in terms of creativity? How to express it is like he can give you data and opinions, but you need to produce your own ideas. (Data and Opinions) 3. So, at this point, will you follow the prediction, or will you follow the research data? Sometimes it does, but usually it is based on one's own preconceptions. (Follow Prediction or Research Data)
Realizable Functions of Artificial Intelligence	Superb Computing Power, Robotic Arm, Facial Recognition, Generative Design, Simultaneous Online Operation, Chip Implantation, Virtual Presentation, Learning Ability, One Click Typesetting, Smart Home, Internet of Things, Automatic Recognition	1. When I used to take a walk in Shenzhen Bay, there were also patrol police cars. They were ridiculously cute and would stop to tell children not to get close to the beach. They even had facial recognition function, which means that if you hurt him or something, he would take your photos and videos. (Facial Recognition) 2. There is a website where you can upload any image and generate any style of past painting experts from it. (Generative Design) 3. Products related to artificial intelligence have now almost realized. For example, there are now robots for delivering packages that can move forward on their own and automatically recognize obstacles, telling pedestrians to move aside. (Automatic Recognition)
User Perspective	Participation Level, Technical Expectations, Interaction Methods, Rhythm, Conflicts between Technology and Nature, Objective Rationality, Standardization,	1. I may pay more attention to my involvement in the design process. I want to present my own thoughts in this product. I must have a part of myself

	High-quality Information, Feedback, Workload, Efficiency, Functionality, Commonality, Suggestions, Logical Processing, Evaluation Criteria, Inspiration, Critical Attitude, Balanced Aesthetics, Fixed Audience, Breakthrough Points for the Audience, Software Operation, Minimal Design, Incomplete Design System, Trial and Error Costs, Hand Drawn Idea Expression, Auxiliary Functions, Sacrifice of Appearance, System Compatibility, Systematic Understanding, Project Progress, Information Processing, Information Volume, Form, Modification, Topic Selection and Creativity, Difficulty in Selection, Learning Cost, Extended Functions, Research and Development, Effective Methods, Preview, Prediction, Knowledge Blind Spots, Paper Feel, High Repeatability, Subjective Sensibility, Creativity, Complicated, Stylish, Emotional, Sophisticated, Ornamental, Minimalist, Aesthetic, Optimal Choice, Respecting	involved in it. (Participation Level) 2. Auxiliary design is quite good. It cannot fully intervene in the design. I will actively embrace innovative technologies, but I am also quite wary of artificial intelligence. I do not have complete trust in artificial intelligence and will consider whether the collaboration between humans and machines can truly achieve the best design results. (Technical Expectations) 3. Sometimes when we use something, we look for the familiar feeling. For example, I used to enjoy drawing on paper. I will rediscover the feeling of drawing on paper on those smart products. So, it is a conflict between technology and nature. (Conflict between Technology and Nature) 4. Artificial intelligence is good at recognizing objective data, but it is difficult to make auxiliary judgments in terms of intuition. Therefore, emotional, or sensory design still needs to be judged by designers or humans. (Objective Rationality) 5. I should still collect it myself, because only I know what information I really need, but it can provide me with higher quality information. (High Quality Information)
Relationship with AI	Copyright, Boundaries, Plagiarism, Authority, Rights, Interests, Privacy, Substitution, Communication, Cooperation,	1. How to define artificial intelligence plagiarism? Is there a standard? I think whether to use it for commercial purposes is a

	Complementarity, Supervision, Vigilance, Betrayal, Friends, Subordinates, Trust, Control, Division of Labor, Auxiliary Tools, High-level Work, Degree of Intervention, Control Ratio, Decision-making, Neuromorphic Intelligence, Forms of Artificial Intelligence, Reality, Virtual Products	standard and requires copyright awareness. (Copyright) 2. Artificial intelligence has fully integrated into design. I will actively embrace innovative technologies, but I am also wary of artificial intelligence. I do not fully trust artificial intelligence. I will think about our cooperation, or in other words, consider the impact of artificial intelligence technology on me. Is it a positive impact? Or is it a negative impact? (Collaboration) 3. Is it difficult for designers to repeatedly modify product models during the development process of both physical and virtual products? Whether it is developing online software products or modeling and rendering physical products, it is difficult to repeatedly modify them. (Virtual Products)
--	---	---

4.2.2.2 Axial Coding

The secondary coding phase of interview transcripts seeks to identify and systematically construct relationships among emergent conceptual categories. Integrate and link open Coding organically—a more comprehensive and clear explanation of the research phenomenon. Axial Coding was built by identifying the main categories of open Coding and connecting various conceptual themes. The data determines the direction in which the design process begins. There are issues with the use of design tools and the development of standards during the design process. From the users' perspective, issues that are overlooked in the design process can be

discovered through research. Therefore, design insights into artificial intelligence needs have been excavated to address user pain points. The logical relationship of the second-level encoding is shown in Figure 4.4.

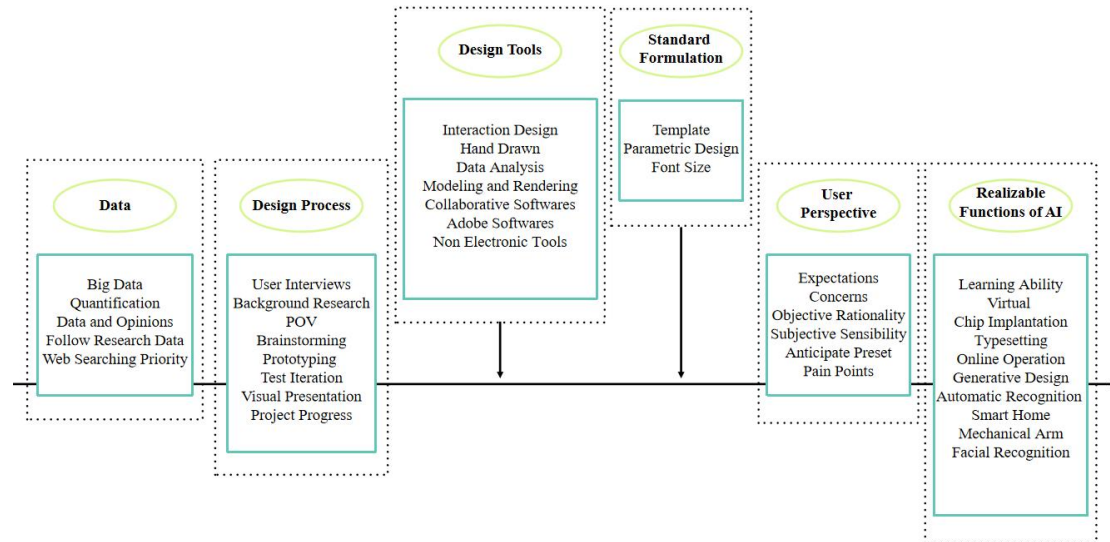


Figure 4.4 The Logical Relationship of Axial Coding

The details of Axial Codes are in Table 4.4.

Table 4.4 Axial Codes

Themes	Main Categories	Connotation	Clusters	Meaning
Improvement of Design Tools	Inconvenient use of design tools - Difficult information processing	The design tools are not intelligent enough to process information, requiring users to spend time collecting and processing information, and collecting and organizing is mostly repetitive work.	Design Tools	Tools required during the design process
	Difficult software	Once the model is determined, it is	Design Tools	Tools required

	operation - Difficult model modification	exceedingly difficult to make further adjustments, especially for more complex models. Can we make it easier to modify the 3D model, including auxiliary design for modeling.		during the design process
	Design tool unable to preview - Prohibitive cost of trial and error	For example, in modeling and rendering software, users cannot preview, or the preview effect is much worse than the actual effect before the modification is completed. This affects users' decision-making during the design process and often discourages them from testing iterations.	Design Tools	Tools required during the design process
Data Collection and Processing	Fixed research population - Difficult to identify design points	The research population is extremely limited, and they do not know how to change the research methods for different target groups, which makes it difficult for users to discover design points and	User Perspective	Examining potential or existing issues in the design process from a user perspective

		increases the blind spots of the research.		
	Knowledge blind spots exist - Lack of design inspiration	When it comes to designs involving other professional knowledge, users are likely to miss useful design solutions due to a lack of understanding of that profession.	Data	Including data sources, data processing, and how to use data in the design process
	Data information collection and processing - Long design cycle	Web searching often do not prioritize information quality, and users need to capture a small amount of valuable information in a huge sea of information, which creates a demand for AI assistance to reduce workload.	Realizable Functions of Artificial Intelligence	Analyzing the existing achievable functions of artificial intelligence from a technical perspective
Design Evaluation and Iteration	Difficult to balance aesthetics - Uncertain rules and norms	Balancing everyone's aesthetics is difficult, and often the final product is vastly different from the original intention. Can we avoid such problems in the preliminary stages.	Standard Formulation	Standards that can be standardized during the design process to save costs
	Conflict between function and form - Difficulty in	Function and form are equally important, sometimes sacrificing the	Standard Formulation	Standards that can be standardized during the design

	Choosing	beauty of form for function, and sometimes having to sacrifice functions for form. Can you provide suggestions in the design process.		process to save costs
--	----------	---	--	-----------------------

4.2.2.3 Selective Coding

Core conceptual categories are identified via a three-level coding process that investigates categorical associations and their functional importance in the interview corpus. Finally, focus on abstracting the core coding. The Selective codes are shown in Table 4.5.

Table 4.5 Selective Codes

Select Core Categories	Representation	Pain Points
Data Collection and Processing	Existence of knowledge blind spots	Lack of design inspiration
	Difficulty in data collection and analysis	Long design cycle, High demand and repetitive workload for AI
	Similar research population	Difficult to identify design points
Improvement of Design Tools	Inconvenient use of design tools	Difficult information processing, Longer design cycle
	Difficult software operation	Difficult model modification
	Design tools is unable to preview	Prohibitive cost of trial and error
Design Evaluation and Iteration	The rules and norms are uncertain	Making it difficult to balance aesthetics
	Conflict between function and form	Difficulty in choosing

This study coded 17 focus group interview data in three stages to summarize the design insights of artificial intelligence-assisted product design and the relationship between codes. This research has found out the role of artificial intelligence in the field of product design. The details are in Figure 4.5.

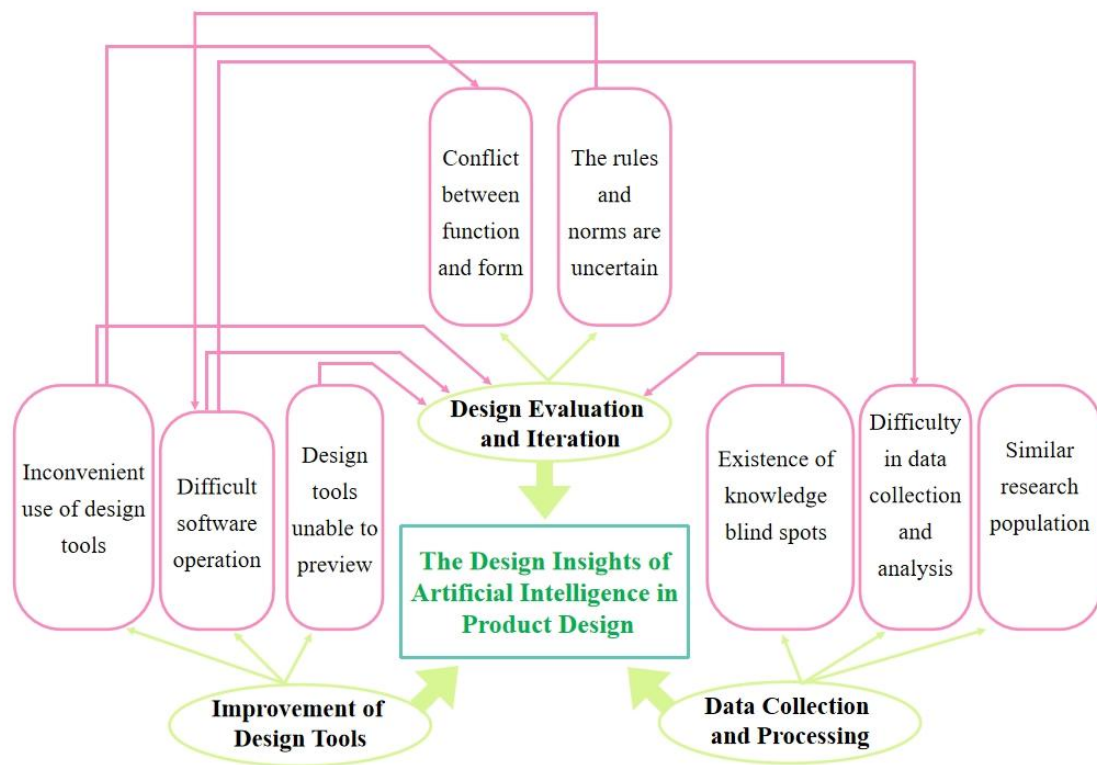


Figure 4.5 The Relationship between Codes

4.2.3 Summary of Results

After analyzing the relevant code, it was found that the three core codes. The selective codes are Data Collection and Processing, Improvement of Design Tools, Design Evaluation, and Iteration. The direct influencing factors and design requirements of artificial intelligence-assisted product design have been identified.

- In terms of data collection and processing, AI can help solve problems such as knowledge blind spots, information collection and processing difficulties, and fixed populations for research. This is consistent with the literature that AI systems can help process substantial amounts of data and provide valuable insights (Feuston & Brubaker, 2021; Vianello et al., 2023).
- In terms of improving design tools, AI can help solve problems such as high software operation difficulty, inconvenient use of design tools, and inability to preview. This is consistent with the literature that AI systems can improve user experience by providing intuitive user interfaces and visualization features (Vianello et al., 2023).
- In design evaluation and iteration, AI can help solve problems such as conflicts between functionality and form and unclear design specifications. This is consistent with the literature that AI systems can provide multidimensional evaluation models and objective evaluation results (Vianello et al., 2023).

However, the limitations of AI in the design process also need to be addressed. For example, overreliance on artificial intelligence may hurt designers' creativity and decision-making abilities (Feuston & Brubaker, 2021). Therefore, when using AI-assisted design, human designers should maintain a dominant position and use AI as an auxiliary tool rather than a decision-maker. Future design practices should pursue a model of human-machine collaboration. While preserving the creativity and insight of

human designers, they should fully utilize AI's data processing and analysis capability (Feuston & Brubaker, 2021) . This collaborative model will help improve efficiency while ensuring the innovation and humanistic value of design solutions.

By making reasonable use of artificial intelligence technology, the user experience of design tools, enhancing the efficiency of data collection and processing, and optimizing the design evaluation and iteration process can be significantly improved. However, the application of artificial intelligence must also be carefully balanced with the protection of human creativity to ensure the overall quality and innovation of the design process.

4.3 Summary and Discussions

AI-assisted design tools are increasingly acknowledged for their potential to transform various industries, such as game design, manufacturing, fashion, and engineering. These tools leverage artificial intelligence algorithms to streamline processes and provide valuable insights (Lankes& Stockl, 2023). In game design, AI-powered tools operate under a mixed-initiative co-creation paradigm, where software and human designers collaborate to suggest and critique design elements, enhancing the overall design process (Lee et al., 2021). Similarly, in manufacturing and industrial design, AI-based tools assist designers in handling complex tasks and optimizing product creation (Gmeiner et al., 2023).

In fashion design, AI tools support designers in creating visually appealing attire and ready-to-wear products that meet design criteria and consumer preferences (Zhang, 2024) . Integrating AI assistance in engineering design,

such as drone design, is essential to manage the increasing complexity of design tasks and understand AI's impact on the design process and effectiveness (Song et al., 2021). These examples highlight the growing role of AI in aiding designers across various domains to overcome challenges and improve their creative output.

Regarding research methodologies, focus group interviews are valuable for conducting constructivist case studies and generating ideas effectively (Moore, 2023). By engaging participants in a group setting, researchers gather diverse perspectives contributing to a comprehensive understanding of the subject. Individual interviews are crucial for collecting in-depth data, making them fundamental in case study research (Moore, 2023). These qualitative methods help explore the requirements and challenges associated with AI-assisted design tools, offering rich data for research.

Furthermore, a detailed exploration of the requirements for AI design tools, particularly in game design, is essential to align these tools with gameplay objectives and enhance the design process (Partlan et al., 2021). By defining specific design-driven requirements, researchers aim to develop tools that cater to designers' unique needs and seamlessly integrate AI into the creative workflow. Understanding these requirements is vital for creating AI tools that effectively support designers in decision-making and creative endeavors.

The collaborative nature of human-AI interaction is a focal point in studies, emphasizing the importance of complementary collaboration between humans and AI systems to boost design efficiency and transcend creative boundaries (Chen, 2024). Analyzing projects involving human-AI

co-innovation highlights AI's potential to enhance human creativity and problem-solving, leading to more innovative design outcomes. This collaborative approach reflects the evolving landscape of design practices, where AI is seen as a strategic partner in driving design innovation and expanding traditional creative processes.

Moreover, AI's role in supporting decision-making processes, such as in clinical imaging, emphasizes the need for thoughtful design considerations governing human-AI interaction (Fogliato et al., 2022). Policies dictating AI inferences' display and sequencing within decision-making workflows ensure AI complements human expertise and contributes meaningfully to decision-making. Studying human-AI workflows' impact on decision-making aims to optimize AI tools' integration in decision support systems and enhance decision quality.

In conclusion, integrating AI-assisted design tools has the potential to revolutionize creative processes across industries. By leveraging AI algorithms, designers can enhance efficiency, creativity, and decision-making, leading to innovative designs. Through a collaborative approach emphasizing human-AI partnership, designers can harness AI's power to push design boundaries and unlock new creative possibilities. Overall, AI has enormous potential in design activities, significantly improving evaluation efficiency and accuracy. Artificial intelligence can improve the efficiency and accuracy of design evaluation in the following aspects.

- Powerful data analysis capabilities: Artificial intelligence can quickly process substantial amounts of data, extract valuable information, and provide an objective basis for design evaluation.
- Quantitative evaluation results: Artificial intelligence can quantitatively score design schemes based on preset evaluation dimensions and standards, making the evaluation results more objective and comparable.
- Multimodal data processing: Artificial intelligence can simultaneously analyze various forms of design data, such as text and images, and comprehensively evaluate all aspects of design solutions.
- Visual presentation: AI can transform complex evaluation results into intuitive visual forms, such as rating radar charts, for designers to understand and improve.
- Simulated expert rating: By learning expert rating criteria, AI can simulate the evaluation process of experts to a certain extent. It can provide professional reference opinions, especially for young designers or scenarios that require extensive design evaluation.

Study 1 explored in depth the actual needs and expectations of designers for AI-assisted tools through qualitative research methods. It provides valuable insights for developing AI tools that better meet the needs of designers. The research results are like the viewpoints of the existing literature on AI-assisted design. For example, the research results support the view that AI systems

can improve user experience by providing intuitive user interfaces and visualization features (Sloterdijk et al., 2017) . Meanwhile, the results also confirm the potential of AI in processing enormous amounts of data and providing valuable insights (Sloterdijk et al., 2017; Yang et al., 2022).

However, this study also has limitations. Firstly, the sample of this study comes from participants with design backgrounds, and most of them are junior designers, which may affect the generalizability of the results. Secondly, the results may have a certain degree of subjectivity due to the use of qualitative research methods. Therefore, the subsequent Study 2 research needs to adopt a mixed study approach, combining quantitative data to validate and expand the findings of this study.

This study provides a new perspective for understanding the role and potential of AI in the design process. It provides a theoretical basis for the future development of AI-assisted design tools. From a practical perspective, the research results can guide the development of AI-assisted design tools in the future, enabling them to meet the actual needs of designers better and, for example, developing AI tools with intuitive interfaces and visualization capabilities to enhance user experience and Developing AI systems capable of processing multimodal data to support more comprehensive design evaluations.

In addition, Study 1 emphasizes the importance of maintaining a balance between human-machine collaboration in AI-assisted design. Future design practices should pursue a human-machine collaboration model, fully leveraging AI's data processing and analysis capabilities while retaining the

creativity and insight of human designers (Gao et al., 2023; Vinnervik, 2022). Although AI has immense potential to improve the efficiency and accuracy of design evaluation, we also need to be aware that AI should serve as an auxiliary tool rather than completely replacing the role of human designers. Designers still need to maintain creative leadership and ultimate decision-making power, using AI as a powerful auxiliary tool to improve work efficiency and design quality (Gao et al., 2023).

Study 1 explored the tool needs of designers in the A era through focus group discussions and grounded theory analysis. These findings offer key insights into understanding the actual needs and expectations of designers and figuring out that the main research direction of this paper is AI-assisted product design evaluation. This research direction directly influenced the questionnaire design for Study 2 and the determination of data requirements. Although this study is limited to a specific group, its results offer important guidance for future research, especially data collection for Study 2. The next chapter will design a specific questionnaire based on these findings.

Chapter 5. Study 2- AI-assisted Design Evaluation Data Requirements

This chapter introduces Study 2, which aims to determine the data types and collection methods needed for AI-assisted design evaluation. The findings of Study 1 showed that design evaluation tasks are an important research area for AI-assisted product design. This study collected designers' understanding of AI tools, usage, and expectations for AI evaluation tools through a questionnaire survey. This study not only confirmed the findings of Study 1 but also offered specific guidance for constructing multimodal product design datasets (Study 3) to ensure that the datasets meet the actual design evaluation needs.

5.1 Introduction

Research on data is crucial for developing artificial intelligence (AI) models that effectively evaluate product design. Consumer reactions and preferences towards AI-designed products are important datasets (Zhang et al., 2022). Understanding consumers' willingness to pay for such products can guide model development by integrating features that meet consumer expectations. The various modes of big data also play a significant role in AI-driven product design (Quan et al., 2023), helping to process, analyze, and utilize information to enhance the design process using AI algorithms. In addition, algorithms related to artificial intelligence technology are also crucial for establishing robust evaluation models (Zhang, 2022). By utilizing artificial intelligence in industrial product design, designers can analyze complex data more effectively, thereby improving decision-making. For example, artificial intelligence is applied to breast cancer screening in health care. This study

also emphasizes the importance of data in ensuring the safety and effectiveness of AI-assisted products (Potnis et al., 2022; Zhong, 2024). Those studies highlight the necessity of high-quality data to train AI models to evaluate product design accurately.

In the context of apply in Improving product design efficiency through g artificial intelligence algorithms to improve product design efficiency, data analysis and professional evaluation are key components (Zhong, 2024). This study indicates that data covering the design process, results, and user feedback is crucial for developing AI evaluation models. In addition, research on product design decision models based on deep learning networks emphasizes the importance of data in improving design decision evaluation (Zong & Wang, 2022) . By adopting a data-driven approach, designers can make informed decisions based on user preferences and design requirements.

The evaluation role of artificial intelligence in visual communication design education has demonstrated its potential (Erturk & Uzumcu, 2022). They are creating AI models through supervised learning using classification datasets to evaluate the feasibility of designing products effectively. The data on intelligent product design in industrial systems such as CNC machine tools provide valuable insights into how artificial intelligence algorithms can optimize product shape and artistic image design (Liu, 2015). These data are crucial for training AI models to make informed design decisions. In addition, personalized data tailored to user needs is crucial for developing customized products (Abbasi& Esmaili, 2024) . Using artificial intelligence algorithms to analyze and interpret customer data, designers can create products that meet specific preferences and requirements. The research on artificial intelligence product design methods based on product semantics also emphasizes the

importance of data (Zhu, 2023). This data-driven approach ensures that the designed product remains competitive in the market.

In short, designing AI evaluation product design models requires a multimodal dataset, including user feedback, industrial design projects, and professional evaluation data. By integrating these data types into the model development process, designers can create AI systems that effectively evaluate and enhance product designs to meet user expectations and market demands. Therefore, this study's purpose and problem setting are highly consistent with the current trend of artificial intelligence applications in product design.

With the diversification and personalization of user needs, traditional product design methods can no longer meet modern market requirements. At the same time, advances in big data and artificial intelligence technology have brought transformative opportunities for product design (Quan et al., 2023). This study aims to determine the data types and collection methods required for AI-assisted design evaluation, which directly addresses key issues in applying artificial intelligence technology in product design. The RQ2 reflects the in-depth thinking of this study on the application of AI technology in product design. This study is consistent with the viewpoint mentioned in the literature that it is necessary to clarify artificial intelligence technology's basic concepts, characteristics, and application modes (Liu & Kim, 2023). By exploring the required data types, Study 2 will contribute to developing a product design evaluation model based on artificial intelligence technology. This research is of crucial importance for the development of product design in the era of artificial intelligence.

5.2.1 Questionnaire Design

This study used structured questionnaires as the primary data collection tool. The design of the questionnaire follows the principles of comprehensiveness, pertinence, and cross-cultural comparison, aiming to explore in depth the needs, attitudes, and expectations of designers towards AI-assisted design evaluation (Wang & Chuang, 2024). The questionnaire structure includes demographic characteristics; design evaluation needs analysis, and personal experience and behavioral influence. This structural design can comprehensively capture the respondents' background information, professional needs, attitudes, and usage of AI technology (Neto et al., 2018).

The questionnaire uses multiple question types, including Single-choice, multiple choice, Likert scale, open-ended, and cross-analysis questions (Neto et al., 2018; Tubadji et al., 2021). Single-choice and multiple-choice questions are used to collect basic information and specific usage. The Likert scale measures respondents' attitudes and opinion intensity towards AI technology. Open questions allow respondents to express their private opinions, while cross-analysis questions help explore the relationship between different variables. This diversified problem-type design helps to obtain comprehensive and in-depth data, providing richer insights for research.

5.2.2 Data Collection

The data collection process adopts the method of online questionnaire survey, using two professional platforms, Sojump and Credamo, to create the questionnaire. The questionnaire is distributed through WeChat and Credamo platforms, inviting designers from various industries to participate.

This method can quickly and widely reach the target group, improving the efficiency of questionnaire collection (Tubadji et al., 2021).

A combination of purposive and snowball sampling in non-probability sampling was adopted in terms of sample selection. The researchers first identified a group of designers who met the criteria and then recommended other eligible participants based on these initial respondents (Neto et al., 2018). This sampling method helps to obtain more relevant samples in specific populations. The combination of the two can construct a more comprehensive technical evaluation framework. American data is to support the external validation of results from different design culture and styles. Chinese designers' data provides insights into local cultural preferences and global design trends, while American data reflects influential international trends that often shape cross-cultural insights. Cross cultural data fusion can enhance compatibility with different design styles. Through comparative analysis, this model can not only accurately meet the needs of Chinese designers, but also absorb international experience, enhance technical stability and cultural inclusiveness.

5.2.3 Data Analysis

This study used a combination of quantitative and qualitative analysis methods. Study 2 focuses on quantitative analysis. Quantitative analysis uses SPSS statistical software, including descriptive statistics, correlation analysis, chi-square test, etc. Descriptive statistics are used to summarize the basic characteristics of a sample and the distribution of variables; Correlation analysis is used to explore the relationship between designers' familiarity with AI tools and other factors; The chi-square test is used to compare

whether the differences between Chinese and American designers in various indicators are statistically significant.

For the answers to open-ended questions, the study used content analysis for qualitative analysis (Tubadji et al., 2021). Researchers extract key themes and patterns by carefully reading and coding respondents' responses to better understand designers' perspectives and needs for AI-assisted design evaluation.

5.3 Results

5.3.1 Sample Characteristics

This study collected a total of 298 questionnaires, and after screening, 297 valid questionnaires were obtained, including 137 questionnaires collected from China and 160 questionnaires collected from the United States. This sample size and distribution provide a good foundation for cross-cultural comparisons. Among them, females account for 44.78%, and males account for 55.22%, with slightly more male participants than female participants. The respondents are concentrated in the age group of 25-34 years old, accounting for 41.08%; Most respondents hold a bachelor's degree or above, accounting for 98.61%; Designers with 1-5 years of work experience account for 43.10%, while those with 6-10 years account for 48.15%. These two groups of designers have the highest proportion. Table 5.1 shows the demographic characteristics of the valid samples.

Table 5.1 Characteristics of the Participants

Name	Options	Region		Total
		China	U.S. A	
Gender	Female	83(62.41%)	50(37.59%)	133

Name	Options	Region		Total
		China	U.S. A	
	Male	54(32.93%)	110(67.07%)	164
Age	18-24	70(95.89%)	3(4.11%)	73
	25-34	41(33.61%)	81(66.39%)	122
	35-44	23(23.71%)	74(76.29%)	97
	45-54	2(50.00%)	2(50.00%)	4
	55-64	1(100.00%)	0(0.00%)	1
level of education	Junior college (or college)	2(50.00%)	2(50.00%)	4
	undergraduate	60(31.09%)	133(68.91%)	193
	master	59(70.24%)	25(29.76%)	84
	PhD or above	16(100.00%)	0(0.00%)	16
Duration of working in the design industry	1-5 years	99(77.34%)	29(22.66%)	128
	6-10 years	21(14.69%)	122(85.31%)	143
	11-20 years	14(63.64%)	8(36.36%)	22
	Above 20 years	3(75.00%)	1(25.00%)	4

5.3.2 Designers' Understanding and Current Usage of AI Tools

The following research findings contribute to understanding designers' perception and usage of AI tools. This study is like the method proposed by Liao et al. (2021) to investigate users' demand for explainable AI (Vianello et al., 2023). Correlation analysis is a quantitative method used to examine the strength and direction of relationships between multiple variables, in order to facilitate understanding of their interrelationships. If the absolute value of the correlation coefficient between two variables is closer to 1, the correlation between the two variables becomes more significant. This study suggests that if the correlation coefficient is below 0.3, there is a weak or no correlation

between two variables; 0.3-0.6 is a moderate strength correlation; A correlation of 0.6 or higher is considered strong. If the correlation coefficient is positive, then the two variables have a positive correlation. Otherwise, there is a negative correlation. The P-value, also known as the significance level, is considered statistically significant if the p-value is less than 0.05, indicating the presence of an effect or difference. If the p-value exceeds the critical value, it is considered statistically insignificant.

A significant positive correlation exists between regions and designers' familiarity with AI ($p < 0.01$, $r = 0.565$). Americans are more familiar with AI tools, as shown in Table 5.2.

Table 5.2 Relationship between Region and AI Familiarity

	Region
Familiarity with the application of AI in design	0.565**

* $p < 0.05$ ** $p < 0.01$

However, gender, age, education level, and length of time in the design industry are not significantly related to familiarity with AI. It is shown in Table 5.3.

Table 5.3 Relationship Between Related Background and Familiarity with AI

	China	U.S. A
Gender	0.162	-0.001
Age	0.005	0.031
level of education	-0.015	0.055
Duration of working in the design industry	0.034	-0.020

* $p < 0.05$ ** $p < 0.01$

Among AI tools, creative tools and image generation tools are highly popular among the designer community. Those are shown in Table 5.4 and Figure 5.1.

Table 5.4 Usage of AI Tools

Topic	Response		Popularity rate (<i>n</i> =297)
	<i>n</i>	Response rate	
Graphic design tools, such as Canva's AI magic design	118	19.97%	39.73%
Creative tools, such as Chart GPT	214	36.21%	72.05%
Image generation tools, such as Midjourney, DALL. E 3	195	32.99%	65.66%
Art generation tools, such as WOMBO Dream	54	9.14%	18.18%
other	10	1.69%	3.37%
Total	591	100%	198.99%

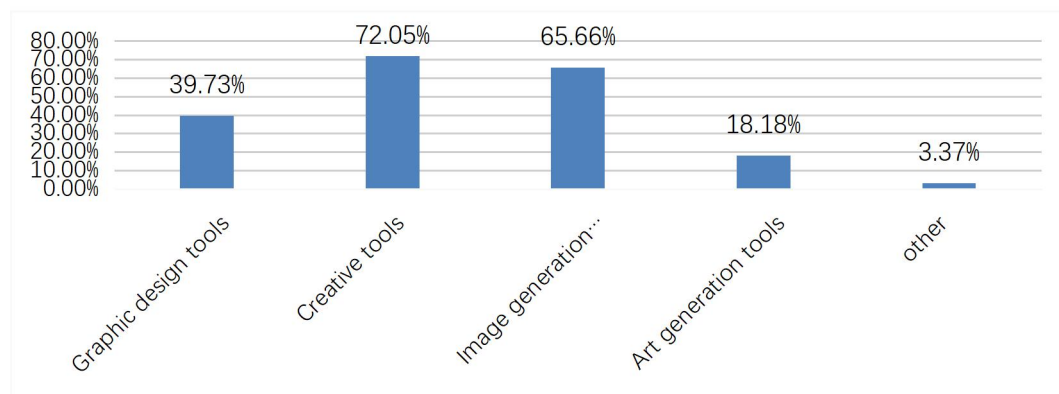


Figure 5.1 Usage of Various AI Tools

55.22% of designers prefer AI as an interactive tool; 31.31% of designers believe that designers and AI should collaborate and participate equally in the design process; 13.47% of designers believe that designers should take the lead in the design process, while AI plays a relatively passive role by

providing information, advice, and data support when needed. Details are shown in Figure 5.2.

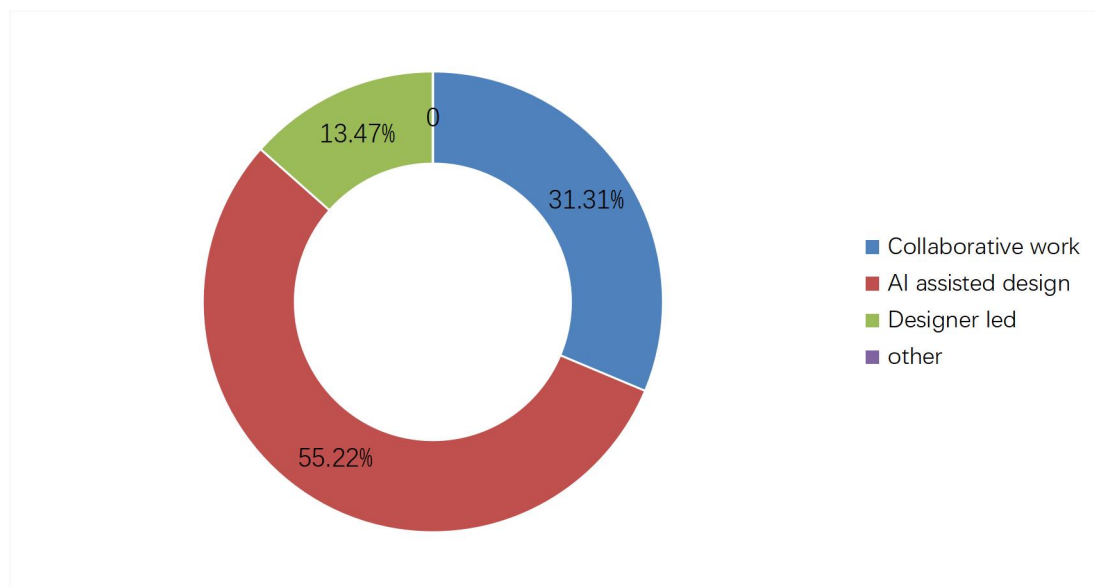


Figure 5.2 The Ideal Interaction between Designers and AI

Designers' demands for AI tools are focused on auxiliary, collaborative, advisory, and maintaining control. These demands guide AI tool designers to consider making AI more adaptable to the designer's work style during the development process while providing efficient, functional, and easily integrated solutions.

As shown in Figure 5.3, among 137 Chinese designers, 121 (88.32%) stated that they had never used AI design evaluation tools, and 16 (11.68%) stated that they had used them. Among 160 American designers, the situation is the opposite, with 20 (12.50%) indicating that they have never used AI design evaluation tools, while 140 (87.50%) indicate that they have used them.

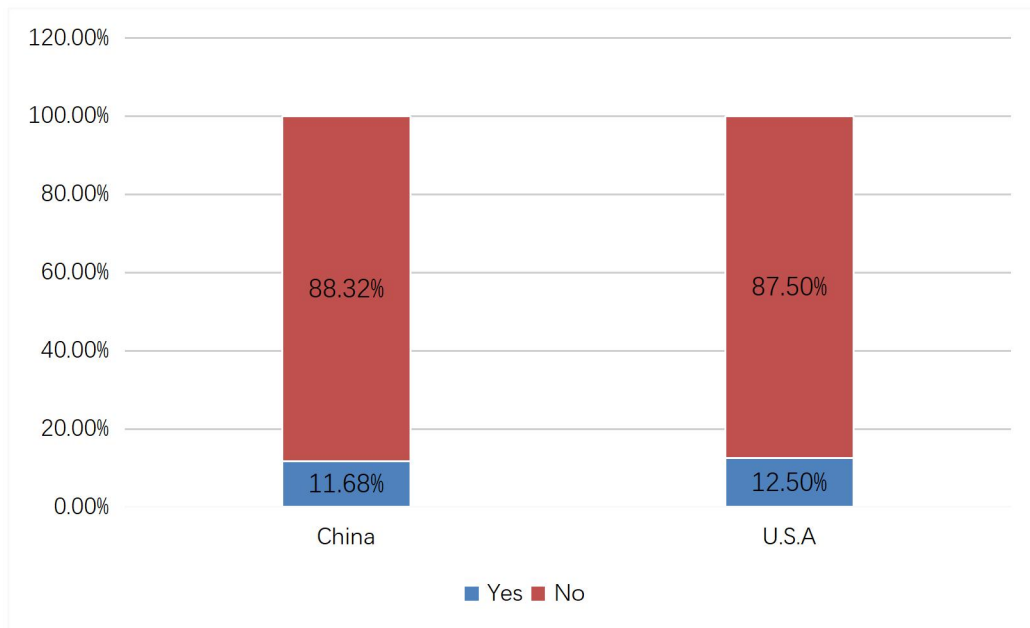


Figure 5.3 AI Technology Evaluation Design Project Usage

According to Table 5.5, there is no significant correlation ($p > 0.05$) between familiarity with AI tools and frequency of use among Chinese designers. This result may be due to the small sample size, as Chinese designers use AI design evaluation tools less frequently. Among American designers, there is a significant positive correlation between familiarity and AI accuracy evaluation ($p < 0.01$, $r = 0.300$). This research result indicates that the higher the familiarity of designers with AI evaluation tools, the higher their evaluation of the accuracy of these tools. A significant positive correlation exists between familiarity and attitude toward AI potential ($p < 0.01$, $r = 0.330$). This indicates that designers more familiar with AI tools have a more optimistic attitude toward the potential of these tools in the future. There is a significant positive correlation ($p < 0.05$, $r = 0.171$) between familiarity and product performance improvement. Although the correlation is weak, it also indicates that designers who are more familiar with AI tools tend to believe that these tools can improve the final product's performance. Designers with experience in using AI are more inclined to accept and use AI tools.

Table 5.5 Relationship between Familiarity with AI Design Evaluation Tools
and Related Indicators

		Familiarity with AI technology-based design evaluation tools	
		China	U.S. A
AI Concept Eval. Frequency	correlation coefficient r	0.487	0.019
	P-value	0.055	0.824
	sample size	16	140
AI Accuracy in Concept Eval.	correlation coefficient r	0.114	0.300**
	P-value	0.675	0.000
	sample size	16	140
AI Creativity in Design Suggest.	correlation coefficient r	0.277	0.050
	P-value	0.299	0.556
	sample size	16	140
AI Potential Attitude	correlation coefficient r	0.277	0.330**
	P-value	0.299	0.000
	sample size	16	140
Adjust Strategy by AI Eval.	correlation coefficient r	-0.231	0.051
	P-value	0.390	0.546
	sample size	16	140
AI Impact on Product Performance	correlation coefficient r	-0.185	0.173*
	P-value	0.494	0.041
	sample size	16	140
* p<0.05 ** p<0.01			

However, various challenges and limiting factors in the design evaluation process may weaken this impact. This leads to a non-significant relationship between familiarity and the willingness to adjust design strategies based on evaluation results, corresponding to the main challenges described by subsequent designers.

Regarding data sources, designers have concerns about the reliability of AI data. The main problem in data analysis is the uncertainty of evaluation criteria. The lack of clear evaluation criteria may make it difficult for designers to trust AI results, affecting strategy adjustments. Regarding resource limitations, AI design evaluation products that designers are familiar with are currently needed.

According to the results in Table 5.6, there are differences in the trust and willingness to use AI design evaluation tools between Chinese and American designers, with American designers showing more obvious trust and positive attitudes.

Table 5.6 Rating of AI Accuracy, Potential Attitude, and Other Aspects by Chinese and American Designers

Title	Name	Region		Total	χ^2	<i>p</i>
		China	U.S. A			
17.AI Accuracy in Concept Design Evaluation	-1.0 Not Very Accurate	2(12.50%)	0(0.00%)	2(1.28%)	77.363	0.000**
	0.0 Moderately Accurate	6(37.50%)	0(0.00%)	6(3.85%)		
	1.0 Quite Accurate	8(50.00%)	76(54.29%)	84(53.85%)		
	2.0 Very	0(0.00%)	64(45.71%)	64(41.03%)		

Title	Name	Region		Total	χ^2	<i>p</i>
		China	U.S. A			
	Accurate					
Total		16	140	156		
18.AI Creativity in Design Suggestions	-1.0 Poor Effect	5(31.25%)	0(0.00%)	5(3.21%)	64.257	0.000**
	0.0 Average Effect	4(25.00%)	5(3.57%)	9(5.77%)		
	1.0 Good Effect	7(43.75%)	56(40.00%)	63(40.38%)		
	2.0 Excellent Effect	0(0.00%)	79(56.43%)	79(50.64%)		
Total		16	140	156		
19.Your Attitude Toward AI's Future Potential in Design	0.0 Neutral	7(43.75%)	2(1.43%)	9(5.77%)	47.459	0.000**
	1.0 Somewhat Optimistic	4(25.00%)	73(52.14%)	77(49.36%)		
	2.0 Very Optimistic	5(31.25%)	65(46.43%)	70(44.87%)		
Total		16	140	156		
20.Adjust Design Strategy Based on AI Evaluation	0.0 Slightly Agree	5(31.25%)	1(0.71%)	6(3.85%)	70.729	0.000**
	1.0 Moderately Agree	8(50.00%)	57(40.71%)	65(41.67%)		
	2.0 Strongly Agree	3(18.75%)	82(58.57%)	85(54.48%)		
Total		16	140	156		
21.AI's Impact on Final Product Performance	Minimal Impact	2(12.50%)	1(0.71%)	3(1.92%)	12.233	0.007**
	Cannot Assess	1(6.25%)	3(2.14%)	4(2.56%)		
	Moderate Improvement	10(62.50%)	90(64.29%)	100(64.10%)		
	Significant Improvement	3(18.75%)	46(32.86%)	49(31.41%)		
Total		16	140	156		

* $p < 0.05$ ** $p < 0.01$

5.3.3 AI Design Evaluation Needs Analysis and Challenges

This part of the research aims to identify the specific needs and challenges designers face when using AI evaluation tools, which is crucial for developing AI systems suitable for designers (Gezici & Tarhan, 2022; Lange, 2024).

43.1% of designers choose to conduct product evaluation during the conceptual design phase. 36.7% of designers choose to conduct design evaluation during the prototype design phase. 1.35% of designers indicate that evaluation will be conducted throughout the design phase. This reflects that designers evaluate in the early and middle stages of the design process to ensure the direction and quality of the design and adjust as necessary, as shown in Figure 5.4.

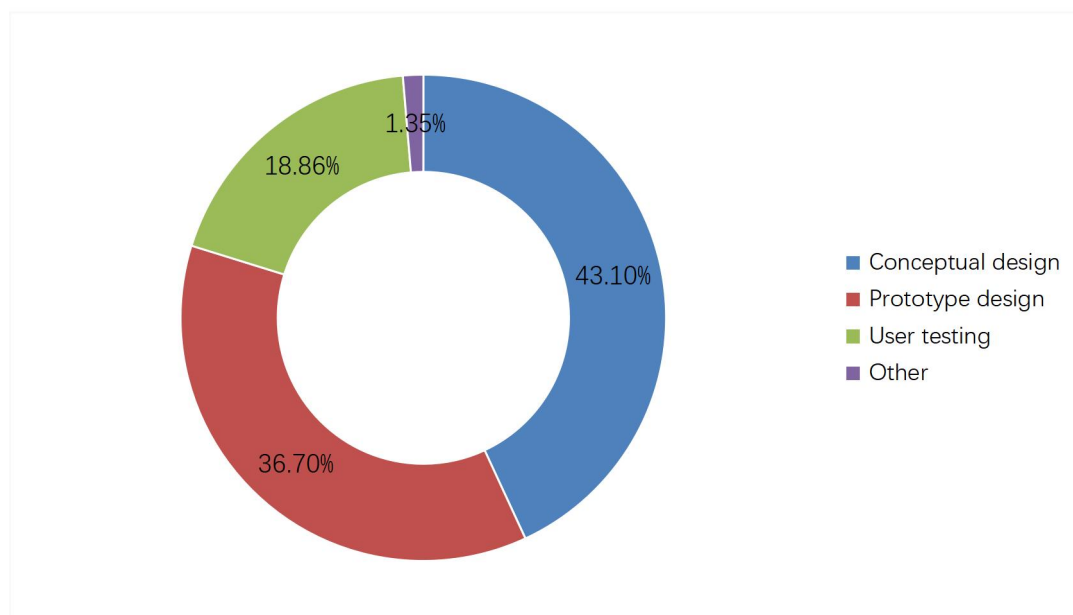


Figure 5.4 The Stage of Conducting Design Evaluation

In the conceptual design evaluation stage, functional implementation has received more attention from Chinese designers. This result may be influenced by the design field in which the sample is located. Designers in

both countries highly value design aesthetics and user experience. The focus on production costs is low among designers in both countries, which may reflect that designers emphasize creativity and user experience more in the conceptual design evaluation process rather than cost-effectiveness.

As shown in Table 5.7 and Figure 5.5, from a statistical perspective, there is no significant correlation between the factors that designers pay attention to and the data sources in conceptual design evaluation ($p=0.680$, $p>0.05$). Among all data sources, the focus on user experience is the highest, especially in user feedback (72.6%), indicating that designers attach significant importance to direct user experience feedback.

Table 5.7 Cross Analysis of Factors of Concern and Data Sources in

Conceptual Design Evaluation

	User feedback	Expert evaluation	Market data	Other
Functional implementation degree	122(55.7)	95(47.7)	88(53.7)	3(100.0)
User experience	159(72.6)	127(63.8)	105(64.0)	3(100.0)
Design aesthetics	139(63.5)	134(67.3)	113(68.9)	2(66.7)
Innovative nature	110(50.2)	105(52.8)	95(57.9)	1(33.3)
production costs	42(19.2)	32(16.1)	43(26.2)	0(0.0)
Other	0(0.0)	0(0.0)	0(0.0)	0(0.0)
Summary	219	199	164	3

chi-square test: $\chi^2=9.272$ $p=0.680$

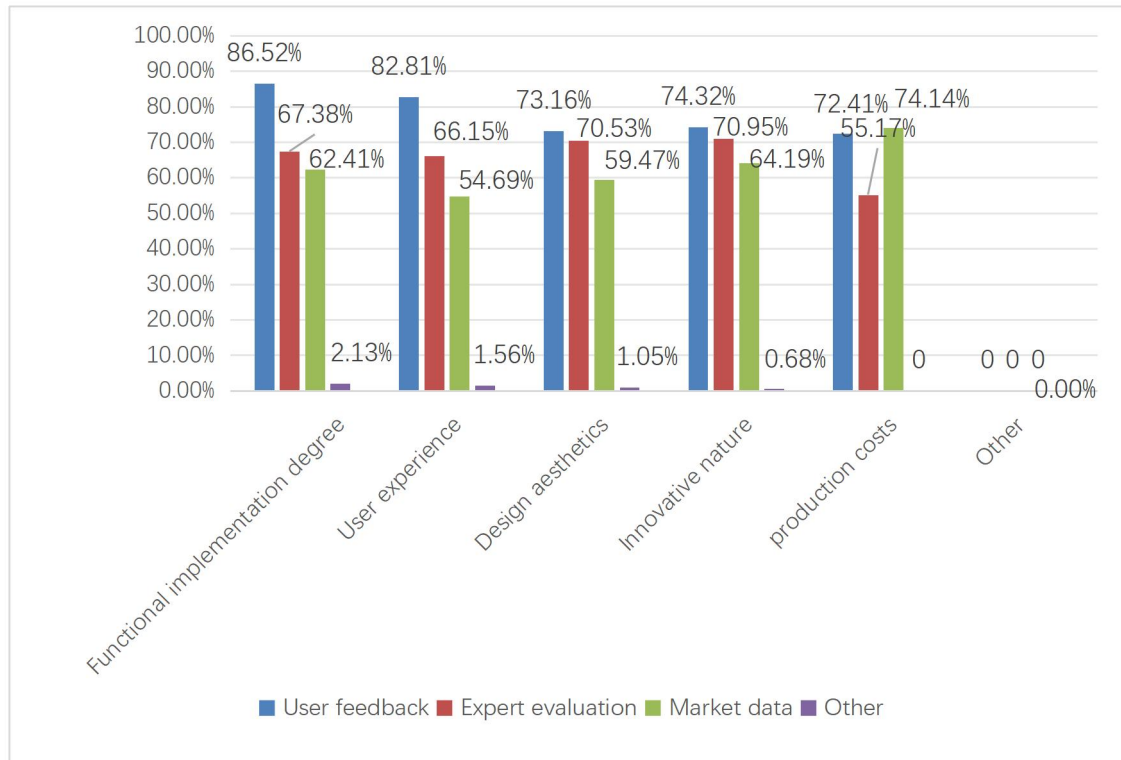


Figure 5.5 Factors of Concern and Data Sources

As shown in Table 5.8 and Figure 5.6, social media, usability testing, and user feedback are the designers' three most mentioned channels when obtaining product design evaluation data. Social media channels have been mentioned 113 times, due to their ability to provide instant and extensive user feedback.

Table 5.8 Data Sources for Design Evaluation

Clusters	Describe	Quantity
A1	social media	113
A2	usability testing	72
A3	user feedback	67
A4	internal feedback	39
A5	post-launch data analysis	22
A6	competitive product analysis	18
A7	design competition	17

establishing evaluation criteria, and ensuring that the analysis results have practical guidance for the design. In the fast-paced design workflow, the limitations of project time, money, technology, and other resources make it difficult for designers to conduct a complete design evaluation.

5.3.4 Expectations for AI Evaluation Tools

These questions explore the specific expectations of designers for AI evaluation tools, including functionality, data types, and presentation of results, which are crucial for developing AI systems that meet user needs (Gezici & Tarhan, 2022; Yüksel et al., 2023).

For users who have not used AI evaluation tools, the reasons for not using AI evaluation tools are shown in Figure 5.8.

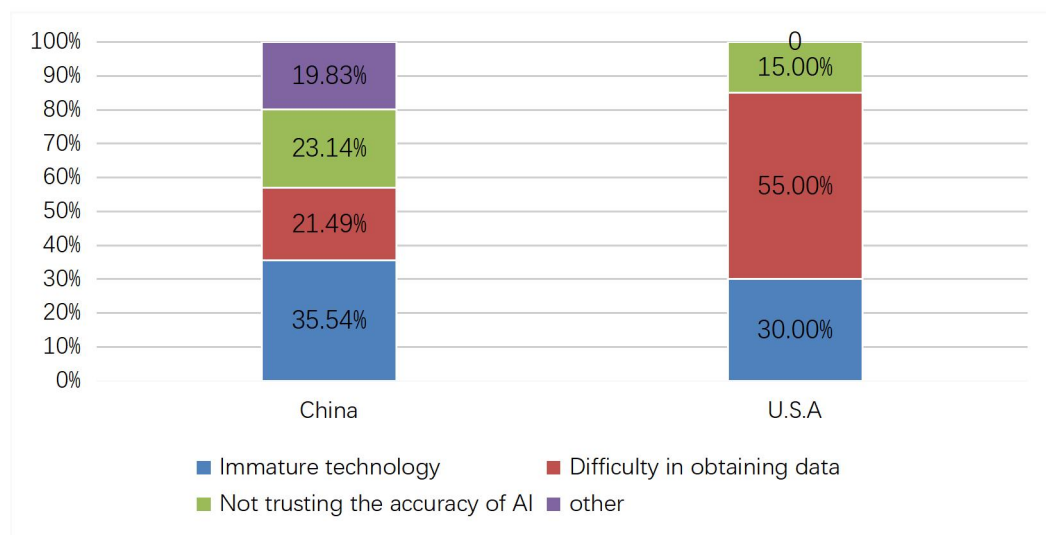


Figure 5.8 Reasons for not Using AI Evaluation Tools

Chinese designers do not use AI tools mainly because they need more trust in AI accuracy and immature technology. Among them, 23 designers chose the "other" option because they were unaware of such products, did not need

them temporarily, and had doubts about the reliability of AI tools, as shown in Figure 5.9. Chinese designers have yet to use AI evaluation tools because of their understanding and market promotion of AI technology. However, the main reason American designers refrain from using AI tools is the difficulty in obtaining data, which may be related to data privacy and access restrictions.

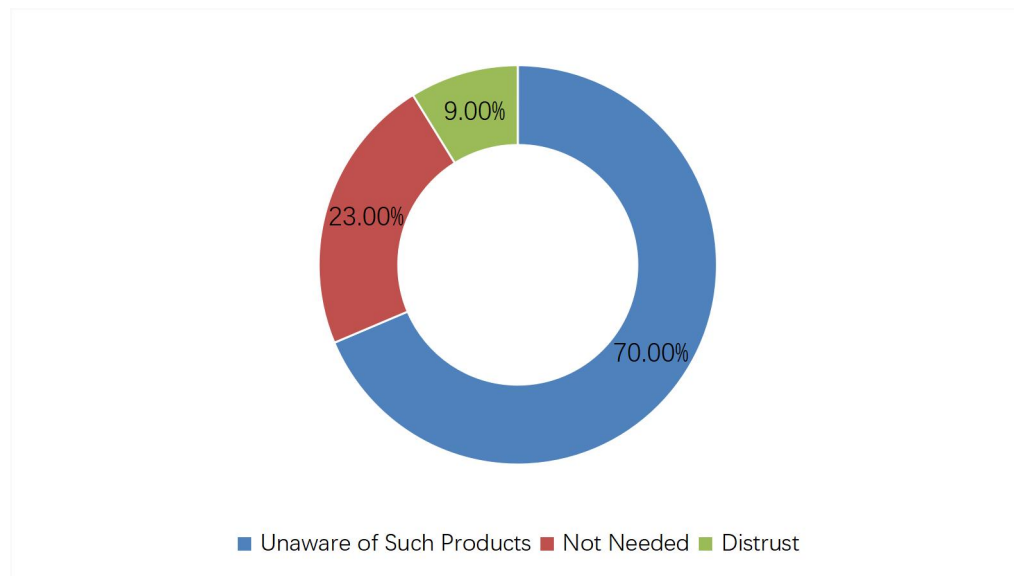


Figure 5.9 Text Analysis of Other Options

As shown in Table 5.10, Chinese and American designers show a certain level of interest and openness in AI evaluation products and their willingness to receive AI technology training or guidance. However, at the same time, they also have a wait-and-see attitude and uncertainty and need more information and practical cases to enhance their confidence and understanding. Figure 5.10 and Figure 5.11 indicate that there is no significant difference between Chinese and American designers in terms of interest in AI evaluation products and willingness to receive AI technology training or guidance. This suggests that designers from both countries have similar attitudes and needs in these areas.

Table 5.10 Interest and Learning Willingness of AI Evaluation

Title	Name	Region		Total	χ^2	p
		China	U.S. A			
23.How interested are you in using artificial intelligence algorithms for product design scoring?	Very Interested, Willing to Try	37(30.58%)	6(30.00%)	43(30.50%)	2.707	0.439
	Interested, Need More Information	61(50.41%)	13(65.00%)	74(52.48%)		
	Uncertain, Need to See Actual Effects	21(17.36%)	1(5.00%)	22(15.60%)		
	Not Interested	2(1.65%)	0(0.00%)	2(1.42%)		
Total		121	20	141		
24.Are you willing to receive training or guidance on the use of AI technology in conceptual design evaluation?	Yes, Very Willing	33(27.27%)	10(50.00%)	43(30.50%)	5.942	0.204
	Interested, Need Further Information	71(58.68%)	10(50.00%)	81(57.45%)		
	Uncertain, Reserved Attitude	15(12.40%)	0(0.00%)	15(10.64%)		
	Not Very Willing, Prefer Traditional Methods	1(0.83%)	0(0.00%)	1(0.71%)		
	Completely Unwilling	1(0.83%)	0(0.00%)	1(0.71%)		
Total		121	20	141		

* $p < 0.05$ ** $p < 0.01$

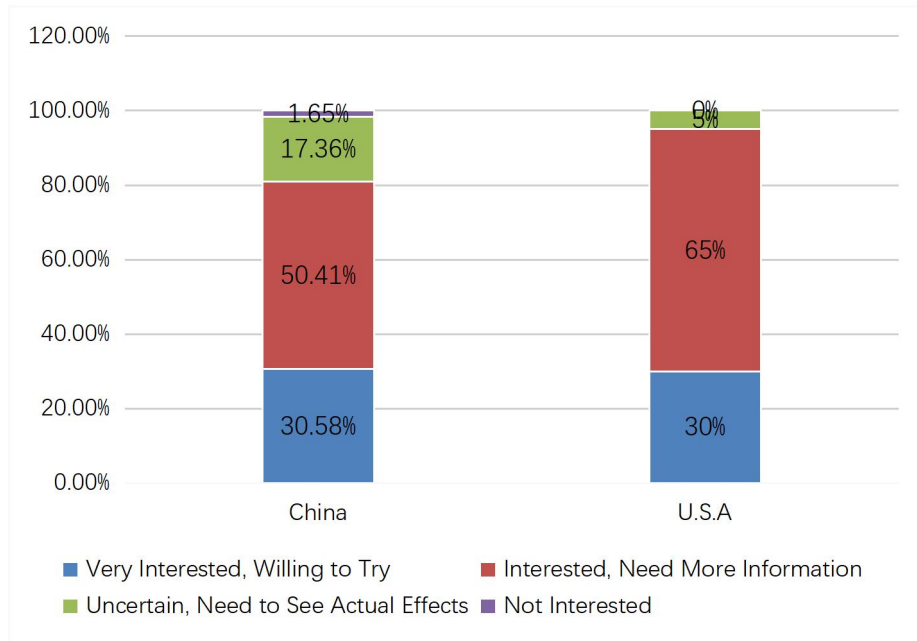


Figure 5.10 The Interest of Designers in AI Evaluation Products

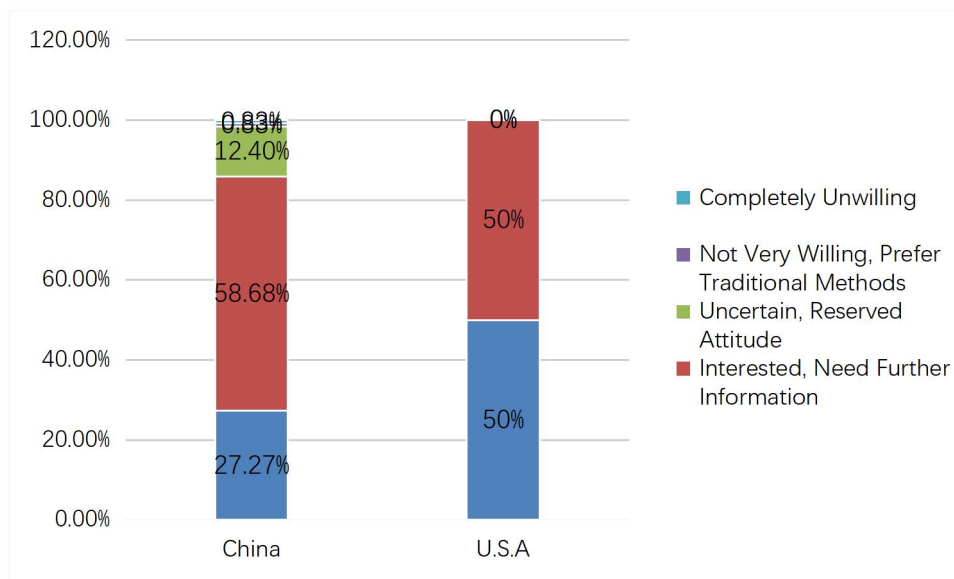


Figure 5.11 The Willingness of Designers to Learn about AI Evaluation Products

Figure 5.12 shows that most users believe AI evaluation has immense potential in the conceptual and prototype design stages.

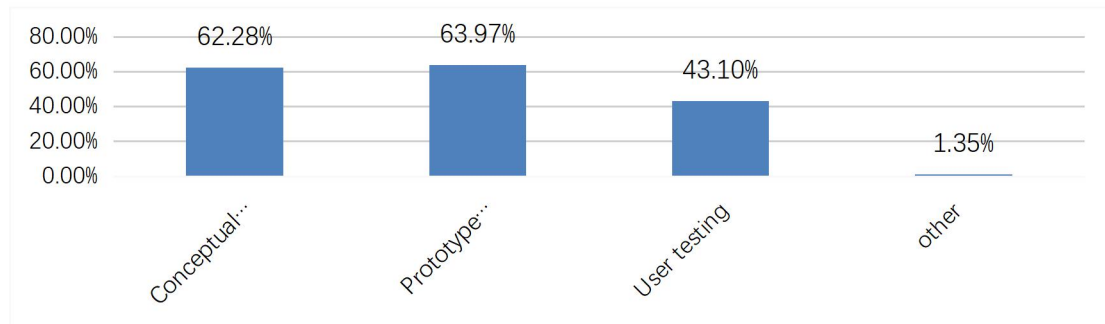


Figure 5.12 Potential Stage of AI Evaluation

Specifically, from Table 11 or Figure 5.13, the advantages that AI is valued at various stages of design evaluation have different emphases. The conceptual design phase emphasizes the discovery of data coverage and design innovation. The prototype design stage places more emphasis on improving efficiency and objectivity. The user testing phase emphasizes objectivity and real-time feedback. This indicates that designers expect AI to provide customized advantages at various design stages to adapt to specific needs at various stages.

Table 5.11 Summary of AI Application Stages and Advantage Crossovers

	Comprehensive Data Coverage	Efficiency in Evaluation	Discovery of New Design Directions and Prediction of Future Performance	Objectivity and Reduction of Human Factors in Evaluation	Real-time Feedback and Suggestions for Improvement	Other
Concept Design	111(72.5%)	105(62.5%)	114(63.7%)	92(63.0%)	43(54.4%)	0(0.0)
Prototype Design	98(64.1%)	131(78.0%)	126(70.4%)	105(71.9%)	54(68.4%)	0(0.0)
User Testing	60(39.2%)	84(50.0%)	88(49.2%)	82(56.2%)	45(57.0%)	0(0.0)
Other	2(1.3%)	2(1.2%)	1(0.6%)	0(0.0%)	1(1.3%)	2(100.0)

	Comprehensive Data Coverage	Efficiency in Evaluation	Discovery of New Design Directions and Prediction of Future Performance	Objectivity and Reduction of Human Factors in Evaluation	Real-time Feedback and Suggestions for Improvement	Other
Total	153	168	179	146	79	2

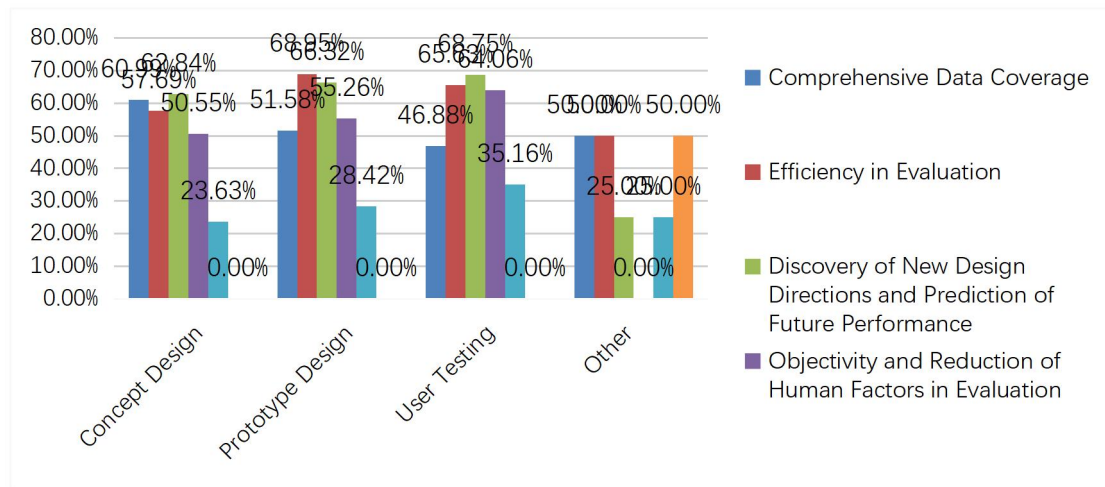


Figure 5.13 Cross Summary of Application Stages and Advantages

As shown in Table 5.12, the standard deviation of each dimension evaluation is between 0.7 and 0.9. This indicates significant differences in the perception of AI technology among the designer community. Overall, the average score of each dimension is >0 , indicating that designers hold a cautious and optimistic attitude toward applying AI technology in design evaluation.

Table 5.12 Designers' Evaluation of AI Technology in Design Evaluation

Title	Minimum value	Maximum value	Mean	SD	median
Reduced evaluation time	-2.000	2.000	1.141	0.814	1.000
Ensure the objectivity of the	-2.000	2.000	1.047	0.895	1.000

Title	Minimum value	Maximum value	Mean	SD	median
evaluation					
Enhance the level of design innovation	-2.000	2.000	0.970	0.967	1.000
Assist in decision-making	-1.000	2.000	1.088	0.757	1.000
Predicting market trends	-2.000	2.000	0.990	0.876	1.000

As shown in Figure 5.14, designers consider user feedback, expert evaluation, and design innovation data to be helpful for conceptual design evaluation.

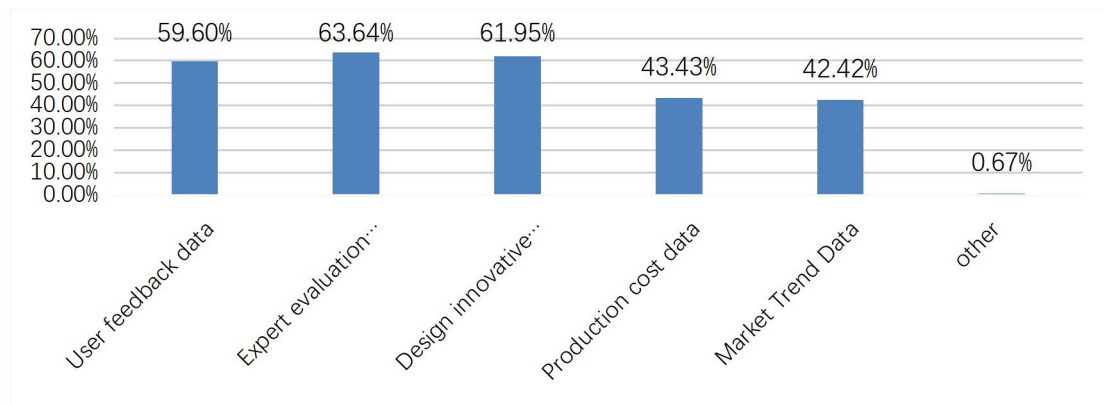


Figure 5.14 Data Types can Assist in Conceptual Design Evaluation

As shown in Table 5.13, the overall response rate reached 271.72%, meaning that each designer selected 2.72 data types on average. This indicates that designers obtain information from multiple perspectives and integrate multiple data types to support their evaluation decisions when conducting design evaluations.

Table 5.13 Summary of Response Rate and Popularity Rate of Data Types

project	Response		Prevalence Rate ($n=297$)
	n	Response Rate	

project	Response		Prevalence Rate ($n=297$)
	n	Response Rate	
User feedback data	177	21.93%	59.60%
Expert evaluation data	189	23.42%	63.64%
Design innovative data	184	22.80%	61.95%
Production cost data	129	15.99%	43.43%
Market Trend Data	126	15.61%	42.42%
other	2	0.25%	0.67%
Summary	807	100%	271.72%

$$\chi^2 = 185.022 \quad p = 0.000$$

As shown in Table 5.14, when evaluating design concepts, there is no difference in each option's co-occurrence, indicating no specific combination preference. Designers tend to use multiple data types comprehensively to ensure the comprehensiveness and accuracy of the evaluation.

Table 5.14 Co-occurrence Matrix of Options of Data Types

	User feedback data	Expert evaluation data	Design innovative data	Production cost data	Market Trend Data
User feedback data	-	-	-	-	-
Expert evaluation data	112	-	-	-	-
Design innovative data	98	112	-	-	-
Production cost data	73	82	96	-	-
Market Trend Data	97	65	84	80	-

As shown in Figure 5.15, designers prefer visual feedback forms of results.

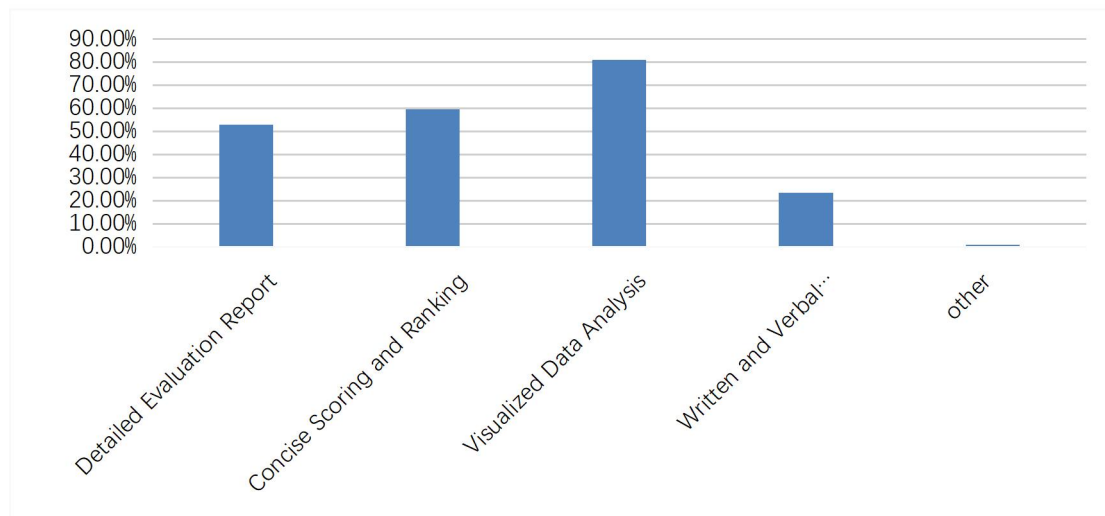


Figure 5.15 1Feedback form of Evaluation Results

As shown in Table 5.15, the overall response rate reached 218.18%, meaning that each designer selected 2.18 feedback forms on average.

Table 5.15 Summary of Response Rate and Popularity Rate of Feedback Forms

Project	Response		Prevalence Rate ($n=297$)
	n	Response Rate	
Detailed Evaluation Report	157	24.23%	52.86%
Concise Scoring and Ranking	177	27.31%	59.60%
Visual Data Analysis	241	37.19%	81.14%
Written and Verbal suggestions	70	10.80%	23.57%
other	3	0.46%	1.01%
Summary	648	100%	218.18%

Note: During the goodness of fit test $\chi^2 = 269.963$ $p = 0.000$

As shown in Table 5.16, in terms of combination preferences, more designers chose concise scores and rankings and visual data analysis.

Table 5.16 Co-occurrence Matrix of Options for Combination Preferences

Project	Detailed Evaluation Report	Concise Scoring and Ranking	Visualized Data Analysis	Written and Verbal Recommendations	Other
Detailed Evaluation Report	-	-	-	-	-
Concise Scoring and Ranking	83	-	-	-	-
Visualized Data Analysis	119	137	-	-	-
Written and Verbal Recommendations	27	48	67	-	-
Other	1	0	1	0	-

The analysis results indicate that designers are cheerful about applying AI in design evaluation, particularly emphasizing its role in assisting, collaborating, and providing suggestions. There is a significant positive correlation between region and designer familiarity with AI, but there is no significant relationship between gender, age, education level, length of time in the design industry, and familiarity with AI. Correlation analysis shows that there is a moderate to strong positive correlation between designers' familiarity with AI and their confidence in its use. The industry experience of designers is positively correlated with their confidence and willingness to use AI, indicating that design experience may promote acceptance of AI. Designers use AI as an auxiliary tool, emphasizing its auxiliary, collaborative, and advisory nature. In addition, designers expect AI tools to provide evaluation recommendations under specific and contextualized constraints to enhance their practical application value in design work.

In summary, designers are cheerful about applying AI in design evaluation, but there are also concerns and expectations. Research can develop AI tools that are more targeted by gaining a deeper understanding of designers' expectations and concerns. Those results have clarified the research direction for subsequent studies regarding data mining, algorithm functionality, and application scenarios. This study aims to promote the research and application of AI technology in the design evaluation field by investigating designers' needs.

5.4 Summary and Discussions

Multidimensional methods are essential to designing AI-assisted product design evaluation tools effectively. Integrating big data and artificial intelligence algorithms is crucial for enhancing the product design process (Quan et al., 2023). Artificial intelligence language models like ChatGPT can influence user experience and design thinking processes and impact product design evaluation (Al-Sa'di & Miller, 2023). Artificial intelligence-driven design tools in mechanical engineering emphasize accelerating development cycles and enhancing product performance features (Al-Sa'di & Miller, 2023). Understanding the collaboration between humans and artificial intelligence is key to designing AI-assisted tools tailored to the needs of different user groups (Dhillon, 2024). Artificial intelligence evaluation data management involves the basic facts of designing and evaluating the performance of artificial intelligence models (Kuo, 2024). Developing automated design systems based on artificial intelligence, such as those for fashion brands, requires a deep understanding of the designer's workflow to optimize the practicality of the tools (Choi, 2023). Other tools, such as computational design, parametric design, and design automation, can significantly enhance the application of artificial intelligence in product development (Manavis,

2023) . Empirical analysis of AI-driven design tools' predictive factors can provide a deeper understanding of the factors that affect tool acceptance and utilization (Chuyen, 2023) . Exploring the metacognitive process in design conceptualization using AI tools can elucidate how AI influences design cognition and decision-making processes (Chang, 2024). Empirical evaluation of artificial intelligence tools is crucial to ensure their effectiveness and reliability in practical applications (Gemrot et al., 2014) . AI design tools for collaborative game creation require collaboration between AI and HCI researchers to effectively collect design heuristic methods (Partlan et al., 2021). The impact of generative design tools on designer behavior highlights challenges such as increased cognitive load and misunderstandings of artificial intelligence recommendations (Saadi & Yang, 2023) . Evaluating the fairness of artificial intelligence systems requires understanding practitioners' processes, challenges, and support needs to design a comprehensive evaluation framework (Madaio et al., 2022).

Study 2 conducted a systematic questionnaire survey to deeply explore designers' needs, attitudes, and expectations toward AI-assisted design evaluation, providing an essential empirical basis for developing AI tools. This research method based on user needs helps ensure that future AI-assisted design evaluation tools can effectively meet the actual needs of designers. The study employed questions and analytical methods, including quantitative and qualitative analysis, providing more comprehensive and in-depth insights. This method can obtain precise statistical results and gain a deeper understanding of the designer's specific ideas and needs. Based on the results of Study 2, this study summarizes the following insights for developing AI-assisted design evaluation tools.

- Artificial intelligence algorithms need to integrate multimodal data. This suggestion is consistent with the viewpoint mentioned in the literature that using big data and AI technology to obtain user needs and expert evaluations more accurately (Quan et al., 2023). This method can indeed provide a more comprehensive perspective for product evaluation.
- The construction of evaluation models requires the application of a deep learning network. The literature mentions the application of deep learning technology in product design, especially its advantages in handling complex design tasks (Quan et al., 2023; Yüksel et al., 2023) . This emphasizes the importance of using advanced AI technology.
- The main implementation scenario of AI evaluation tools is conceptual design. The study also found that designers believe that the conceptual design and prototype design stages are the most useful periods for AI evaluation tools (Rath et al., 2022). This provides important guidance for applying AI tools in the product design process.
- The evaluation results of intelligent models need to be visualized through feedback. The research results and related studies indicate that designers prefer to obtain visual feedback on results (Rath et al., 2022).

However, this study also has limitations. For example, the samples in this study were only from the United States and China. The research results may not directly apply to groups from other cultural backgrounds. In short, designing AI-assisted product design evaluation tools requires a comprehensive approach that combines big data analysis, artificial

intelligence algorithms, user experience evaluation, and empirical evaluation. By using multimodal data sources and algorithms, the development of subsequent artificial intelligence product design evaluation models can be guided by evidence-based practice. This is to ensure the effectiveness of this model in the product design evaluation process.

Study 2 determined the data types and collection methods required for AI-assisted design evaluation through a questionnaire survey. This study not only verified the findings of Study 1 but also offered specific guidance for constructing the multimodal product design datasets (Study 3). Although the sample may have limitations, the results provide key information for the next stage of dataset construction to ensure that the datasets meet the actual design evaluation needs. This lays the foundation for constructing multimodal datasets in the next chapter.

Chapter 6. Study 3 Part I- Construction of Multimodal Product Design

Datasets

This chapter details the construction process of the multimodal product design datasets. Based on the data demand analysis of Study 2, conceptual product design samples holding multimodal data, such as text and images, were collected and processed. This study conducted a preliminary analysis of the datasets through natural language processing and other technologies and provided basic data support for developing the next AI evaluation models. The construction of these datasets directly supported implementing the AI evaluation model in Study 3.

6.1 Introduction

Based on the data requirements analysis results of Study 2, it is indicated that subsequent research needs to collect conceptual product design evaluation samples containing multimodal data such as text and images. It is the key foundation for implementing the artificial intelligence evaluation model in the next Chapter 7.

In artificial intelligence (AI) and deep learning, integrating multimodal datasets has become an important research area. Multimodal datasets contain various kinds of data, such as text, images, videos, and audio, providing a rich source of information for AI systems (Zhou et al., 2020). Using AI to understand and reason about multimodal data is essential for improving AI

capabilities (Chen, 2021) . AI systems can achieve a more comprehensive understanding and analysis by integrating multiple data types, including visual, auditory, and textual information (Zhou et al., 2020).

The fusion of multimodal data involves converting data from various sources into a unified representation, thereby enhancing the overall performance of the AI system (Hao et al., 2022) . This fusion process aims to create a compact representation that captures complementary information from each modality (Li et al., 2022). Utilizing multimodal fusion methods can improve the performance of various applications, such as urban traffic scene analysis and dangerous driving behavior identification (Gao et al., 2021; Ran et al., 2022).

In product design and conceptualization, AI-assisted design processes benefit from multimodal datasets. By combining AI techniques such as natural language processing and deep learning, designers can leverage the power of AI to analyze and interpret multimodal data for innovative product design solutions. Using AI in conceptual design tasks can enable AI systems to effectively process and respond to multimodal queries, thereby streamlining the design process.

In summary, integrating multimodal datasets in the AI-driven design process can potentially transform all areas of product design. Researchers and practitioners can unlock new possibilities in data-driven design and conceptualization by leveraging the power of AI, deep learning, and natural language processing. This study conducts a multimodal product design evaluation study in this context, laying the foundation for integrating

multimodal datasets with AI technology. This chapter is significant for advancing the AI-assisted product design decision-making process.

6.2 Methods

Based on the questionnaire survey results in Chapter 5, this study collects concept design evaluation data to support the subsequent development of a multimodal evaluation model. In addition, this study also uses natural language processing technology to intuitively present the relationship between text data and scores. This analysis method can provide accurate statistical results and comprehensive and in-depth guidance for the development of AI-assisted design evaluation tools.

6.2.1 Datasets Collection

The questionnaire survey results show that to better establish an AI product design evaluation system at this stage, collecting a large amount of manually evaluated product design data for algorithm learning is necessary. Therefore, in the initial data collection stage, 1,354 design projects of undergraduate students majoring in product design were collected. These design projects have 4,905 conceptual design renderings, and the primary data examples are shown in Figure 6.1. The design works include student personal information, such as name and student number, and work information, such as work name, design description, and work score. Eight product design lecturers or professors score the work scores, and the evaluation results are in percentage.



Figure 6.1 An Example of the Basic Data

These basic data provide the possibility for model training. However, its label needs to be more singular; there are no more relevant weights to reflect, and it is prone to overfitting. Therefore, to train the model better, collecting more quantitative product design evaluation data with labels and weights is necessary. This study thanks the Shenzhen Industrial Design Industry Association (SIDA) for providing a useful design platform. The researcher contacted the organizing committee of the Goldreed Industrial Design Award competition. Through communication, they also face evaluation and screening work in the preliminary evaluation stage of their works, and they have a strong demand for multimodal design evaluation. They are willing to provide competition data for this study. Therefore, a large product design database is authorized and collected to develop artificial intelligence models.

6.2.2 Datasets Processing Methods

The datasets have eight categories of works, containing a total of 7921 design projects. The project categories include transportation, industrial equipment, information processing, home life, sports and health, public facilities, cultural and creative, and communication design. Among them, one competition work consists of pictures and design descriptions. In the initial evaluation stage,

each project is rated by five experts based on design, technology, marketing, investment, and media indicators. The total score of the initial evaluation is obtained by a weighted average of the scores given by five experts.

Pytorch is used to build a model frame. Numpy is an introductory Python package for processing numbers. PIL is used to process pictures. Sklearn is used for machine learning, and pandas are also an essential library for processing data. The dataset was separated into three parts: the training set, validation set, and test set, and the ratio is about 98:1:1. This is an effective way to prevent this model from overfitting (Figure 6.2).

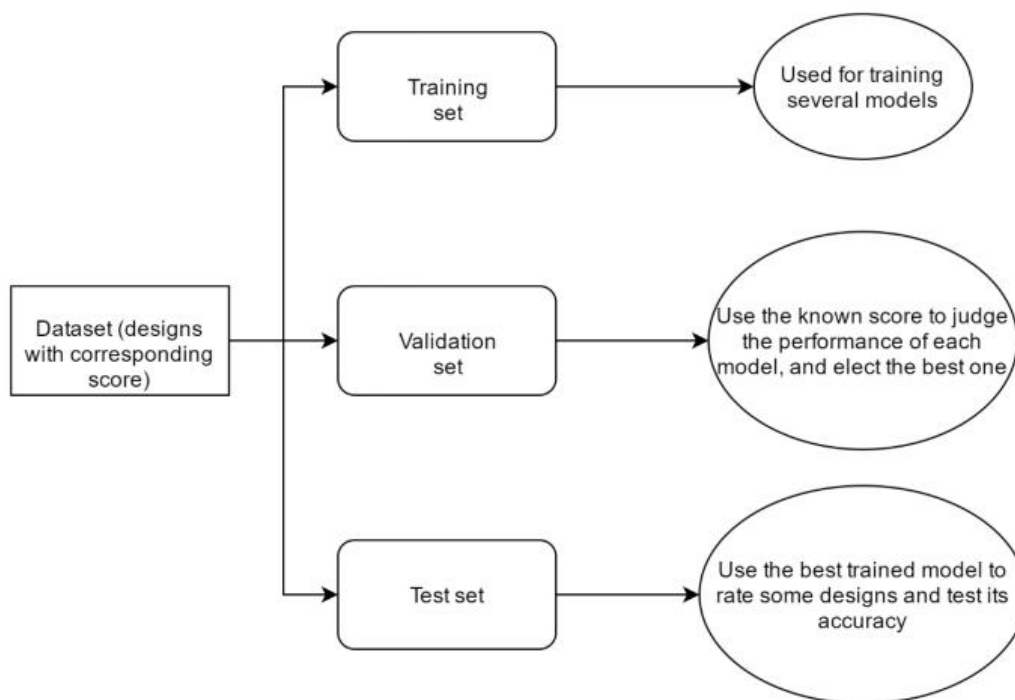


Figure 6.2 Data Classification

Then, the training set was preprocessed to make it more suitable for the training model and reduce irrelative affected elements, like graying, resizing, and augmentation of the data. Graying can reduce computational complexity.

Adjusting the size can help reduce errors in images of varied sizes. Sharpen and enhance edge information. Smoothing is used to reduce noise in images. Increase sample richness by flipping and rotating. These methods can make the original image more efficient and suitable for deep learning, as shown in Figure 6.3.

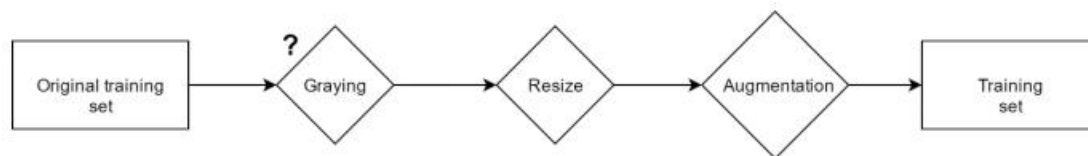


Figure 6.3 Data Processing

6.2.3 NLP Text Analysis

This study introduced natural language processing technology to analyze the collected data and explore the relationship between text and ratings. This exploration reflects the mining of multimodal data in this study. This study conducted a word cloud analysis on the collected data set and introduced a deep learning method to analyze the rating data. The specific analysis content will be presented in the results section. That section analyzes the length and content of the design descriptions of works with different rating levels and explores the relationship between text data and ratings.

6.3 Results

6.3.1 Database Statistical Analysis

In this study, 67,000 tagged digital image files were collected, totaling 77 G. The database relies on the product design competition to collect 8,024 pieces of global design works. The design projects are divided into nine categories:

998 pieces of industrial equipment, 465 pieces of communication design, 706 pieces of public facilities, 895 pieces of household products, 2465 pieces of concept products, 456 pieces of transportation, 1091 pieces of cultural creativity, 283 pieces of information products, and 665 pieces of sports and health products. The details are in Table 6.1.

Table 6.1 Projects Information

Category	Quantity
Industrial Equipment	998
Communication Design	465
Public Facilities	706
Household Products	895
Concept Products	2465
Transportation	456
Cultural Creativity	1091
Information	283
Sports and Health	665
Total	8,024

The evaluation results are evaluated by more than 50 design industry experts from all over the world. Every project has an evaluation label by five experts. The five review experts come from design, technology, marketing, investment, and media. The final label of each project is the average of the evaluation results of these five experts. Functionality, innovation, user-friendliness, aesthetics, and sustainability are the five criteria that are convenient for evaluating designs, each weighing 20%. The details are in Table 6.2.

Table 6.2 Evaluations Information

Expert Field	Evaluation Criteria	Weight
<i>Design</i>	Functionality	20%

<i>Technology</i>	Innovation	20%
<i>Marketing</i>	User-Friendly	20%
<i>Investment</i>	Aesthetics	20%
<i>Media</i>	Sustainability	20%

There is an example of the database (Figure 6.4). Due to copyright issues, the design work was blurred, but the relevant information of the dataset was displayed.

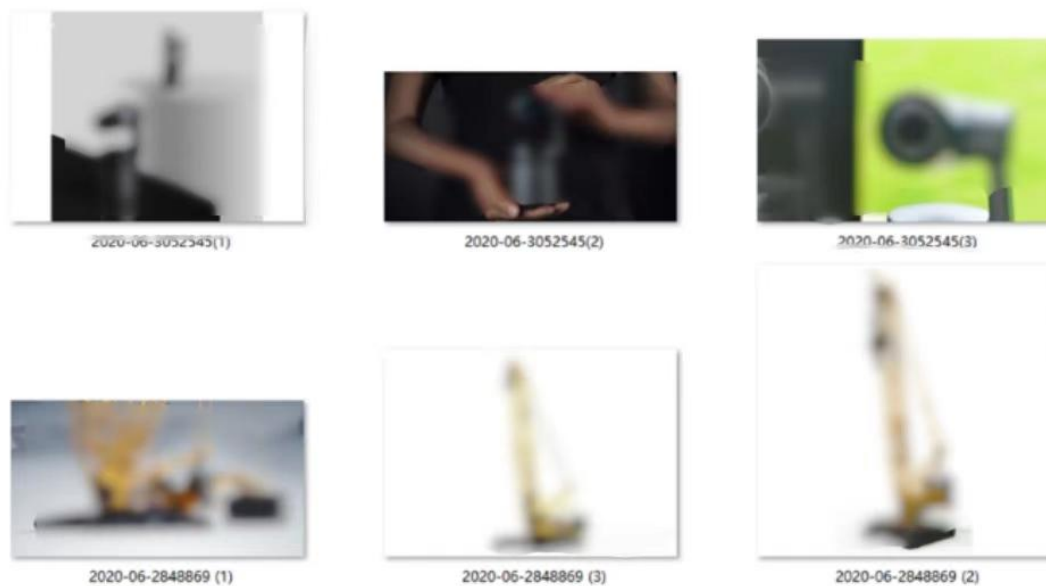


Figure 6.4 An Example of the Database

Figure 6.5 shows the distribution of scores for eight categories on the dataset. The horizontal axis represents categories. The vertical axis represents the score. Assorted colors represent distinct categories. The graph shows high and low-scoring product projects, with most projects receiving moderate scores. The overall distribution of rating results is quite reasonable.

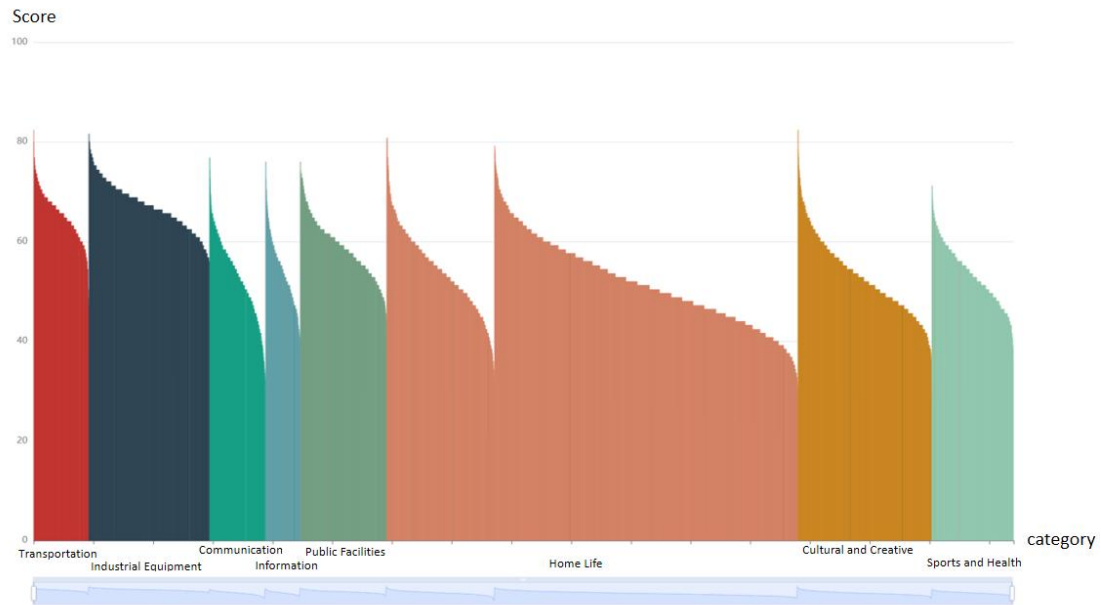


Figure 6.5 The Distribution of Scores

This study selected the ratings of 10 experts with ratings. Divide the score range from 0 to 100 into 20 intervals with 5 points to calculate the score. The distribution shape of the ten expert ratings is like a Gaussian distribution, and there will be no bimodal extreme ratings. Moreover, there are few high and low-scoring works. Therefore, the ratings of the ten experts are reasonable. Figure 6.6 shows the evaluation results of one of the experts.

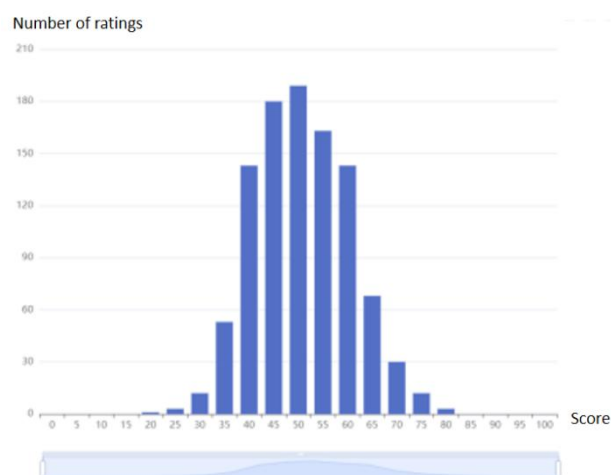


Figure 6.6 Example of Experts Rating Results

The following is a matrix chart of correlation coefficients for design, technology, market, investment, and media. Figure 6.7 is the Spearman's rank correlation coefficient (SRCC) correlation coefficient matrix. Spearman's ρ constitutes a nonparametric statistical technique that evaluates the ordinal association between variables by measuring their ranked correlation. This method determines the degree to which a monotonic relationship characterizes the interdependence of two ranked datasets.

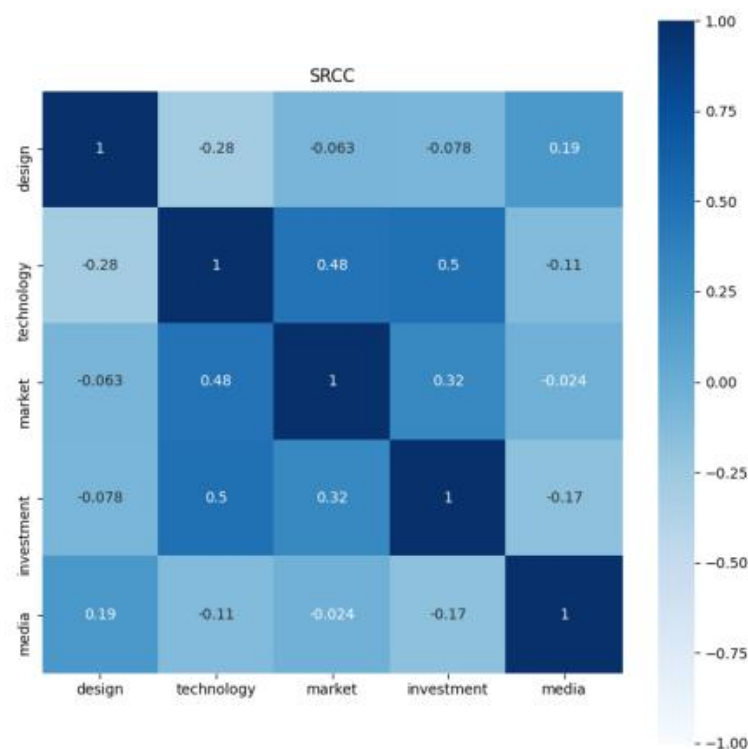


Figure 6.7 SRCC

Figure 6.8 shows the Pearson linear correlation coefficient (PLCC) correlation coefficient matrix. The Pearson correlation coefficient measures the linear relationship between two datasets. Like other correlation coefficients, this one varies between -1 and +1, with 0 implying no correlation. Correlations of -1 or

+1 imply an exact linear relationship. + 1 and -1 represent positive and negative correlations, respectively.

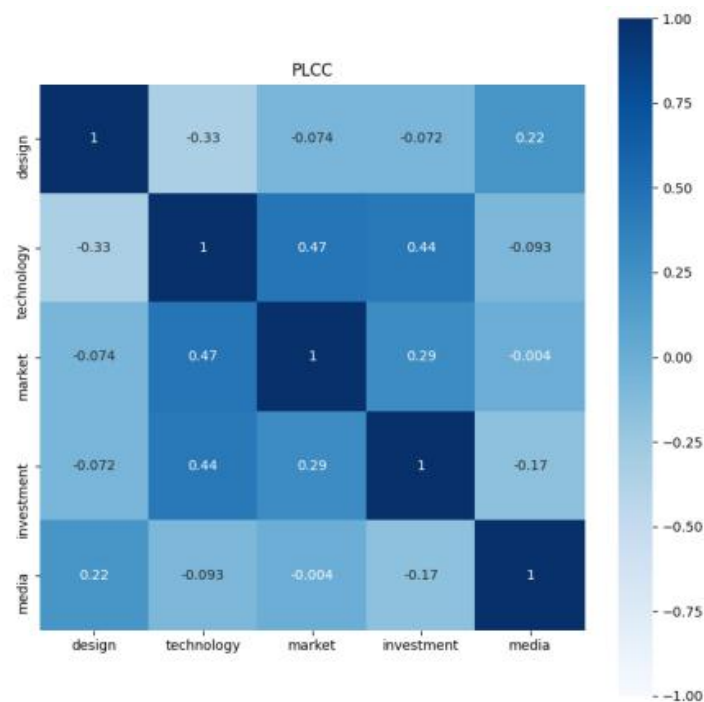


Figure 6.8 PLCC

The correlation coefficient matrix results show that the correlation coefficients are concentrated in the range of $[-0.2, 0.5]$, and there is no absolute correlation between the scores of the five scoring indicators. This indicates that the situation where one indicator scores high and another indicator scores high is not absolute.

6.3.2 Data Analysis by Natural Language Processing

Study 3 also conducted NLP analysis on the dataset, focusing on the Study 3 also conducted NLP analysis on the dataset, focusing on the relationship between text data (design specifications) and ratings. Data analysis found that the richness of content and depth of vocabulary expression in design

specifications can affect expert ratings. The dataset is divided into three categories based on score ranges: low score (10-40 points), median score (40-70 points), and high score (40-70 points). Three categories of high-frequency vocabulary are presented in the form of word clouds.

The low score word cloud is shown in Figure 6.9. This word cloud contains basic product design elements like water, chairs, bottles, etc. These words reflect the fundamental concepts in design but need more innovation and complexity. This is consistent with the characteristics of low-scoring works, which typically only contain basic design elements and functional descriptions.



Figure 6.9 The Word Cloud of the Low Score

Figure 6.10 shows the median score of the word cloud. This word cloud highlights words such as space, system, and life, indicating that design considers a broader range of systems and user experiences. At the same time, words such as development and experience appear in the figure, reflecting a certain degree of innovative thinking. This aligns with the characteristics of medium-quality works, which focus on user needs and innovation but may need to be deeper.



The most prominent words in the high-score word cloud are equipment, design, and technology, which also include words such as efficiency, process, and development, as shown in Figure 6.11. This reflects high-level design thinking, integrating advanced technology, innovative processes, and systematic considerations. High scoring works typically display more complex design concepts, innovative solutions, and a deep understanding of user needs.



This classification method is consistent with the scoring criteria used in research, and high-quality design works typically demonstrate higher levels

of innovation, functionality, and user consideration. Meanwhile, this analysis method also reflects the effectiveness of using keywords and cluster analysis in design evaluation.

In addition, this study also analyzed the relationship between text length and score, as shown in Figure 6.12. The horizontal axis represents scores, and the vertical axis represents text length.

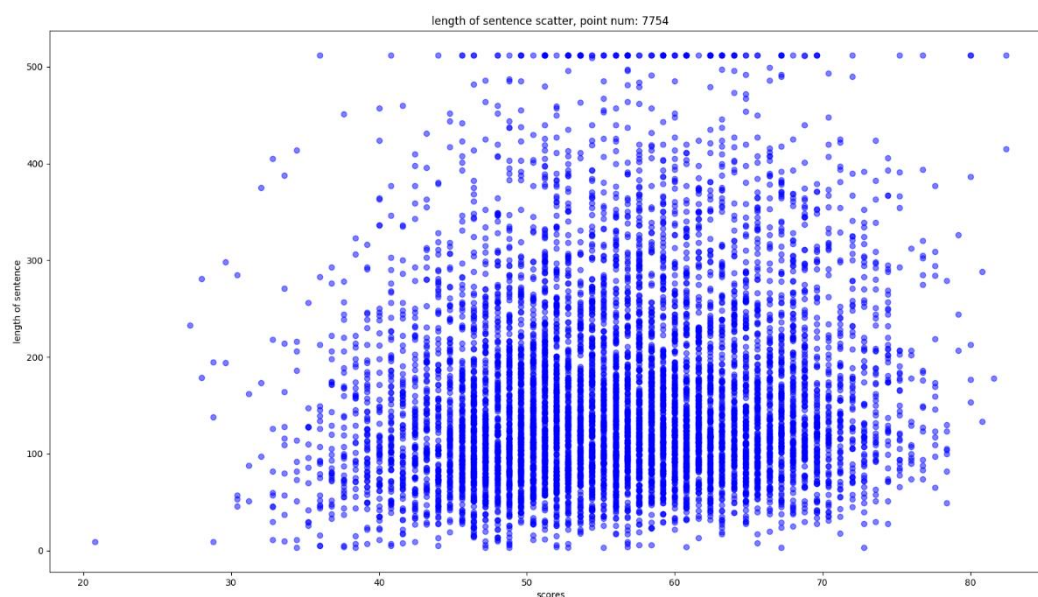


Figure 6.12 The Relationship between Text Length and Score

This study reveals the complexity and multidimensional characteristics of product design evaluation by analyzing the relationship between length and rating. The scatter plot of 7754 data points is shown in the figure, with the horizontal axis representing the design score and the vertical axis representing the length of the design description. From an overall trend perspective, as ratings increase, the length of design specifications shows an increasing trend, indicating that more detailed and comprehensive explanations often accompany high-quality designs.

However, it should be noted that the length of the design specifications is not the only determining factor for the rating. As shown in the figure, there is a wide overlap in length distribution in each rating interval, which means that other factors, such as visual effects, innovation, and functionality of the design, also play a significant role in the rating. This complex evaluation mechanism reflects the demand for multidimensional evaluation in modern product design, including but not limited to aspects such as product appearance, user experience, technological innovation, and sustainability.

6.3.3 Functional Contribution Analysis of Multimodal Data

The multimodal data for product design in this study mainly includes image data and text data. The image data mainly includes design renderings, design illustrations, and so on. Text data mainly consists of design specifications, including design concepts, functional descriptions, and so on.

Image data mainly serves the following three functions in the model. Firstly, visual feature extraction is used to extract visual features such as shape, color, and material of the product through CNN network. Next is aesthetic evaluation, which evaluates the aesthetic value and visual appeal of products based on visual features. Finally, there is functional identification, which identifies the possible functional attributes of the product through its appearance.

The contribution of textual data to the model includes the following three aspects. Firstly, it is about understanding the design intent, by analyzing the design explanatory text, to comprehend the designer's creative intention. Next

is the addition of functional descriptions, where textual descriptions supplement functional information that images cannot fully express. Finally, there is concept positioning analysis, which extracts design concept positioning related information from product description text.

Collaboration of different modal data is achieved through the following methods. Firstly, feature complementarity, where image features and text features complement each other, providing more comprehensive product information. Next is cross validation, which involves cross validation between different modal data to improve the reliability of the evaluation. The ultimate goal is fusion optimization, which achieves optimal fusion of multimodal features through attention mechanisms and other methods.

6.4 Summary and Discussions

This chapter details the construction process of the multimodal product design dataset, which is the key foundation for implementing artificial intelligence evaluation models. By integrating text and image data, this study can capture all aspects of product design more comprehensively, providing solid data support for subsequent AI-assisted design evaluation. During the review process, experts consider the visual presentation of the design and deeper factors such as design concepts and material selection. These concepts and considerations are usually reflected in the design description, so high-scoring designs often have longer descriptions. As a written expression of design thinking, the design description provides an important window for reviewers to understand design intent and innovation.

Based on these data analysis results, future research directions should focus on developing a more comprehensive evaluation model. The model needs to integrate multimodal data, including text description, visual design, material information, etc., to reflect the overall situation of product design. Using multimodal algorithms to train the evaluation model will help to capture all aspects of the design more accurately and provide more comprehensive and objective evaluation results. This approach not only improves the accuracy of the evaluation but also provides designers with more valuable feedback, promotes product innovation, and improves design quality.

The natural language processing technology and deep learning methods used in this chapter, especially in text data analysis, demonstrate the potential of AI technology in design evaluation (Saidani et al., 2021; Singh et al., 2024). This method can extract key information from design descriptions and reveal the potential relationship between text features and design scores.

However, this study also recognizes that there are limitations in the dataset construction process. For example, the representativeness and diversity of the data may need to be further improved to ensure that the AI model can adapt to various kinds of product designs (Tsang & Lee, 2022; Yüksel et al., 2023). In addition, optimizing data preprocessing methods is also an important direction for future research. Therefore, with relevant literature, future research from the following points can be conducted.

- Further expand the data source, including more diverse product categories and design styles, to improve the representativeness of the dataset (Yüksel et al., 2023).

- Explore more advanced multimodal fusion techniques to integrate text and image information better and improve AI models' understanding and analysis capabilities (Saidani et al., 2021).
- Develop more sophisticated natural language processing models to capture subtle semantic differences and innovative concepts in design descriptions (Saidani et al., 2021; Singh et al., 2024).
- Consider introducing factors such as sustainability and eco-design to make the dataset better reflect current design trends and social needs (Saidani et al., 2021).

In general, the multimodal product design dataset constructed in this chapter lays the foundation for AI-assisted design evaluation, but there is still room for improvement. Future research should focus on optimizing data collection and processing methods and how to more effectively use AI technology to extract design features, thereby promoting the application of AI in product design and innovation (Tsang & Lee, 2022; Yüksel et al., 2023).

This chapter details the construction process of the multimodal product design dataset based on the data requirements analysis results of Study 2. By collecting and processing product design samples holding multimodal data such as text and images, this study provides basic data support for developing AI evaluation models. Although the dataset may have specific limitations, it provides the necessary data foundation for implementing the AI

evaluation model in Study 3 and directly supports the next stage of model development. The next chapter will develop an AI evaluation model based on these datasets.

Chapter 7. Study 3 Part II- Implementation of AI Evaluation Model

Chapter 7 is the core part of Study 3. Based on the multimodal product design dataset constructed in the previous chapter, Study 3 is committed to realizing an artificial intelligence evaluation model. This chapter will explore how to use deep learning algorithms, especially multimodal algorithms that combine computer vision and natural language processing technologies, to train AI evaluation models. By comprehensively using methods such as visual feature extraction, text analysis, and multimodal fusion, this Study aims to build an intelligent model that can comprehensively evaluate product design. Implementing this AI model is a continuation and sublimation of the previous data collection work and a solid technical foundation for subsequent human-computer comparison experiments. The research results of this chapter will directly affect the performance of the AI model in actual evaluation tasks, thereby verifying its potential and practicality in assisting design decisions.

7.1 Introduction

In the field of artificial intelligence (AI) for product design evaluation tasks, the selection of algorithm models plays a significant role in determining the effectiveness and efficiency of the evaluation process. Studies provide insights into various algorithmic models and methods that can be used for evaluating tasks. For example, in medicine, studies use deep neural networks to evaluate medical image diagnostic analysis, emphasizing the importance of AI research design in evaluation (Kim et al., 2019; Park & Han, 2018).

In addition, Zong and Wang (2022) proposed a product design decision model based on deep residual networks aimed at improving the efficiency of design evaluation while reducing the impact of decision preferences on the design process. This model demonstrates the application of artificial intelligence in simplifying design decisions. Mondal's (2024) research elucidates how computational intelligence technologies, including evolutionary computing methods, connectionist systems, and reward based learning paradigms, can promote the development of mechanical design optimization processes, highlighting their transformative potential in design evaluation methods.

Quan et al. (2023) reviewed research on product design driven by big data and artificial intelligence. This study emphasizes the role of artificial intelligence algorithms in processing and analyzing various data modalities to facilitate product design. This highlights the synergistic effect of big data analysis and artificial intelligence in optimizing the design process. Related to this, Verganti et al. (2020) explored how artificial intelligence can fundamentally change innovation and design processes by improving scalability, expanding boundaries, and enhancing adaptability, emphasizing the transformative impact of AI on traditional design practices.

In addition, Jha et al. (2022) outlined the best practices for evaluating artificial intelligence algorithms in their research on nuclear medicine and artificial intelligence, emphasizing a comprehensive framework for assessing technology efficacy, clinical decision-making, and post-deployment efficacy. This structured approach can adapt to AI-driven product design evaluation tasks to ensure thorough evaluation. Zhang (2022) explored the integration of AI technology in industrial product design, emphasizing the collaborative

relationship between AI and design innovation in creating comprehensive design solutions.

In summary, when considering the most suitable algorithm to evaluate artificial intelligence product design, Study 3 used empirical evaluation research methods. The following research requires the use of deep neural networks to optimize decision models, utilizing big data analysis and metaheuristic algorithms for efficient evaluation. This Study integrates best practices in industrial design to achieve a comprehensive approach to product design evaluation driven by artificial intelligence. This Study ensures that the subsequent evaluation prototypes of artificial intelligence product design achieve optimal results and higher efficiency in the auxiliary design process.

7.2 Methods

Figure 7.1 shows the architecture of a multimodal deep-learning model for product design evaluation. This model combines text and image features to provide a comprehensive product prediction evaluation. It includes the following key components: text feature extraction, image feature extraction, feature fusion, multilayer perceptron (MLP), and multi-dimensional evaluation.

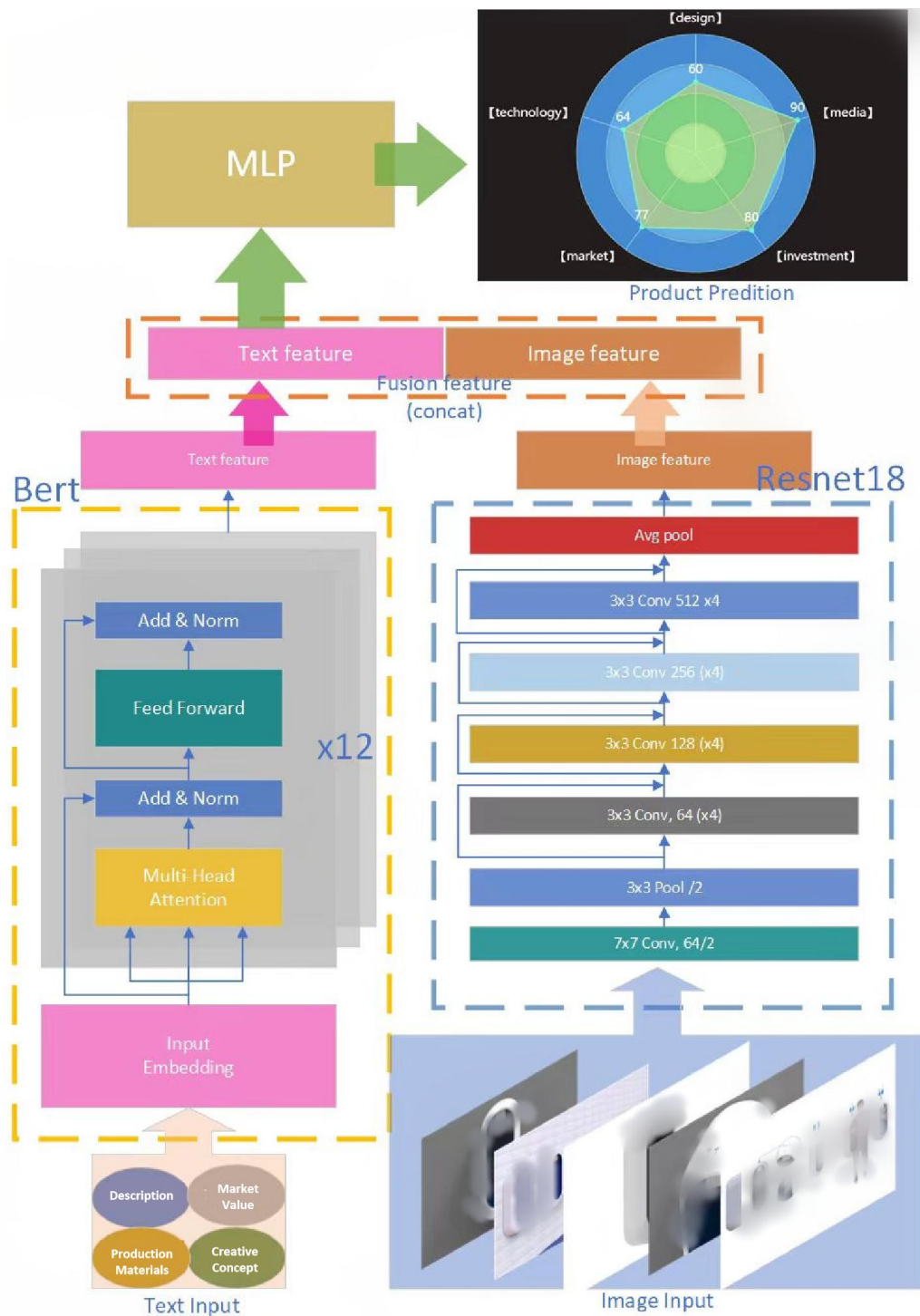


Figure 7.1 The Architecture of the Multimodal Deep Learning Model

Text feature extraction uses the BERT model to process text input, including information such as product description, market value, production materials, and creative concepts. Image feature extraction uses the ResNet18

convolutional neural network to process image input. Feature fusion combines text and image features through the concatenation method to form a comprehensive feature representation. The multilayer perceptron (MLP) uses MLP to process and classify the fused features further and output the final product prediction results. The model output includes scores for multiple dimensions, such as design, technology, market, investment, and media, intuitively displayed as radar charts. This multimodal approach fully uses natural language processing and computer vision technology. It can comprehensively analyze product design from both text description and visual presentation perspectives, providing an innovative and intelligent solution for product evaluation.

7.2.1 AI Evaluation Algorithm

Based on previous literature, this Study compares different algorithms. Multi-algorithms can be divided into two categories: CV and NLP. The following text will discuss these two types of algorithms separately. The visual algorithms employed were Vgg16 (Pal et al., 2023), ResNet18 (He et al., 2016) , DenseNet121 (Bulbul et al., 2022) , CustomNet (Madni, 2024) , and Vision Transformer (ViT) (Dosovitskiy, 2020) . The language understanding capabilities of the BERT model (Devlin, 2018) were leveraged to interpret and analyze the textual data accompanying the product designs.

7.2.1.1 Visual Feature Extraction Module

For extracting Visual Features, Study 3 exploited off-the-shelf deep learning networks as Visual Feature Extraction, like Vgg (Pal et al., 2023), ResNet18 (He et al., 2016) , Densenet (Bulbul et al., 2022) , Vision Transformer (ViT) (Dosovitskiy, 2020) , and a customized neural network named CustomNet.

The size of input images and details of these model pipelines will be presented as follows: Table 7.1, Table 7.2, Table 7.3, Table 7.4, and Table 7.5.

Table 7.1 Vgg16

Layer info	Layers Specification	Output size
Block 1	conv, 3x3, stride 1, 64x2 Max pool, 2x2, stride 2	[112, 112, 64]
Block 2	conv, 3x3, stride 1, 128x2 Max pool, 2x2, stride 2	[56, 56, 128]
Block 3	conv, 3x3, stride 1, 256x3 Max pool, 2x2, stride 2	[28, 28, 256]
Block 4	conv, 3x3, stride 1, 512x3 Max pool, 2x2, stride 2	[14, 14, 512]
Block 5	conv, 3x3, stride 1, 512x3 Max pool, 2x2, stride 2	[7, 7, 512]
Dense	Max pool fcx3	[1000]

Table 7.2 ResNet18

Layer info	Layers Specification	Output size
Block 1	conv, 7x7, stride 2, 64x2 Max pool, 3x3, stride 2	[112, 112, 64]
Block 2	conv, 3x3, 64x2	[56, 56, 64]
Block 3	conv, 3x3, 128x2	[28, 28, 128]
Block 4	conv, 3x3, 256x2	[14, 14, 256]
Block 5	conv, 3x3, 512x2	[7, 7, 512]
average pool	average pool 1x1x512	7x7 average pool
fully connected	512 x 1000 fully connections	1000

Table 7.3 Desene121

Layer info	Layers Specification	Output size
Block 1	conv, 7x7, stride 2 Max pool, 3x3, stride 2	[112, 112]
Dense Block 1	conv, 1x1, x6 conv 3x3, x6	[56, 56]
Transition Layer 1	conv, 1x1 average pool, 2x2, stride 2	[56, 56] [28, 28]
Dense Block 2	Conv, 1x1, x12 conv 3x3, x12	[28, 28]
Transition Layer 2	conv, 1x1 average pool, 2x2, stride 2	[28, 28] [14, 14]
Dense Block 3	conv, 1x1, x24 conv 3x3, x24	[14, 14]
Transition Layer 3	conv, 1x1 average pool, 2x2, stride 2	[14, 14] [7, 7]
Dense Block 4	conv, 1x1, x16 conv 3x3, x16	[7, 7]
Final Layer	average pool, 7x7	[1, 1]
	fully connections	1000

Table 7.4 CustomNet

Layer info	Layers Specification	Output size
Block 1	conv, 3x3, stride 2, 64x2 Max pool, 3x3, stride 2	[112, 112, 64] [56, 56, 64]
Block 2	conv, 3x3, 64x2 Max pool, 3x3, stride 2	[28, 28, 64] [14, 14, 64]
Block 3	conv, 3x3, 128x2 Max pool, 3x3, stride 2	[14, 14, 64] [7, 7, 64]
average pool	average pool 1x1x512	7x7 average pool
fully connected 1	64 x 128 fully connections	128

fully connected 2	128 x 256 fully connections	256
-------------------	-----------------------------	-----

Table 7.5 ViT

Layer info	Layers Specification	Output size
Linear Projection of Flattened Patches	Linear Projection	[196, 16, 16]
Transformer Block	Transformer x 12	[196, 768]
fully connected 1	fully connections	128

7.2.1.2 Vision-based Product Rating Algorithm

Given a group of product design photos, where N is the number of photos, this Study could get a visual feature group, where N is the number of photos with the Visual Feature Extraction Module accounted for in the previous section. Due to the uncertainty of design photos, we attempt to unify the data format of extracted features with the element average operator; then, the averaged vector implicitly representing the Design Product comes out.

$$v_i = f_v(x_i), \#(1)$$

$$av = \frac{v_1 + v_2 + \dots, v_N}{N}, \#(2)$$

This vector is followed by a two-layer Multilayer Perceptron (MLP) for the purpose of feature reshaping and producing Product Rate. The first perceptron layer could reduce the last dimension of the vector to 768, and the second perceptron layer could fit this dimension-unified vector into the Product Rate. It is a 5-dimension vector of technology, design, media, investment, and market. The whole procedure is shown in Figure 7.2.

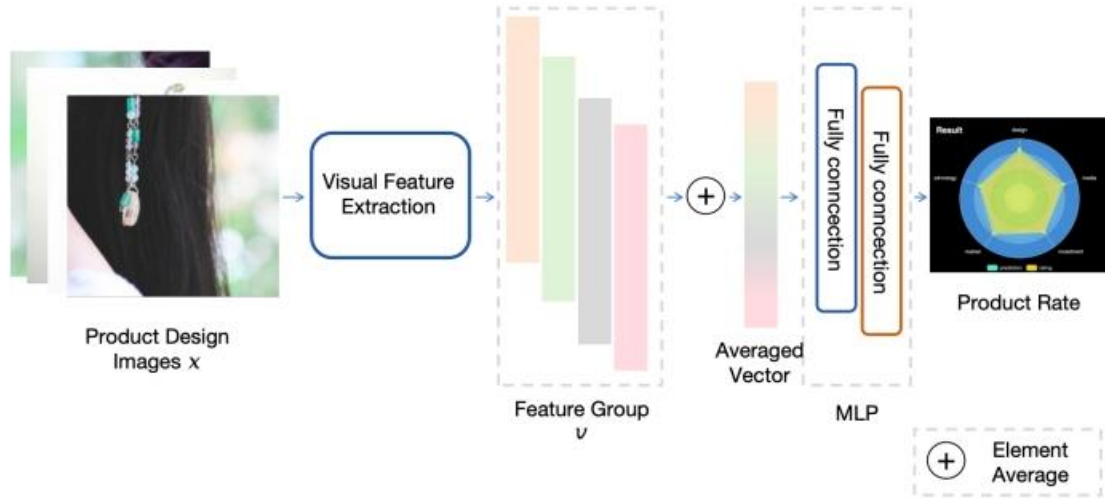


Figure 7.2 Vision-based Product Rating Algorithm

7.2.1.3 Bert-based Product Rating Algorithm

Based on data analysis, text information associated with product design could provide vital references for Product Rating. The BERT, a pre-trained transformer model for language understanding, is used to extract text features. Specifically, this step only considers the relevant text as the text data input, like product description, product value, craft, and creativity. This Study uses an off-the-shelf workpiece tokenizer from hugging face to separate these relevant sentences into individual tokens. Then, three sequential embedding operators follow with these tokens: token embeddings, segment label embeddings, and word position embeddings. These embedding features are taken as the input of 12 transformer blocks.

Therefore, a concatenated feature symbolizing text information of the product is mined. It employs MLP to reshape features and produce product Rates. The output dimension of each layer is described in the same way as in this section. The whole procedure is shown in Figure 7.3.

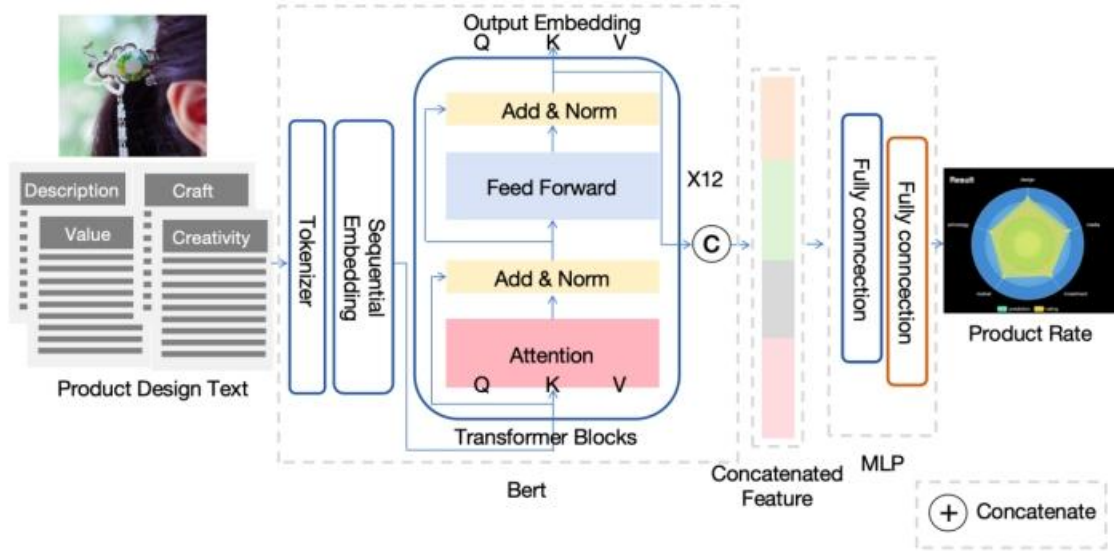


Figure 7.3 Bert-based Product Rating Algorithm

In practice, the same settings in "Attention Is All You Need" for these special models were used, as shown in Figure 7.3 above. This Study uses sine and cosine functions with different frequencies as the positional embedding.

$$PE(pos, 2i) = \sin(pos/10000^{2i/d_{model}}), \# (3)$$

$$PE(pos, 2i+1) = \cos(pos/10000^{2i/d_{model}}), \# (4)$$

The i is the dimension. The pos is the position of the sentence. And the d_{model} is the last dimension for feature vector which is 768.

In the Attention module, the sequential embedding will be computed as query matrix Q , key matrix K and value matrix V . Then, it could get the output of the attention module as follows. The d_k is the dimension of the key matrix.

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \# (5)$$

The feed-forward component employs a densely connected neural architecture comprising sequential linear transformations, incorporating rectified linear unit (ReLU) nonlinearities between layers.

$$FF(x) = \max(0, xW_1 + b_1)W_2 + b_2 \#(6)$$

The x is the output of $\text{LayerNorm}(x + \text{Sublayer}(x))$ and the $\text{Sublayer}(x)$ is the function implemented by the sub-layer. The W_1, b_1, W_2, b_2 are learnable parameters. The neural architecture maintains consistent 512-dimensional representations at both input and output layers, while employing an expanded 2048-dimensional hidden representation in its intermediate processing stage.

7.2.1.4 Vision-Language Fusion for Product Rating Algorithm

Multimodal combination models were also evaluated. These included integrations of Vgg16 & BERT and ResNet18 & BERT (Shao et al., 2021). These combination models aim to merge the insights from visual and textual analysis, providing a holistic view of the product designs that mirror how humans perceive and evaluate products.

For better performance of Product Rating, this Study exploited multimodal features to assess a piece of design product. In the process of visual feature extraction, the same structure was used in the Vision-based Product Rating Algorithm section. The same structure was used in the text feature extraction process's Bert-based Product Rating Algorithm section. However, there is a full connection with Text Feature for the purpose of aggregating features into

a unified feature shape, which is 768 data dimensions. Sequentially, this Study concatenated the extracted visual feature and text feature together, and a two-layer multilayer perceptron (MLP) was followed to fit the output dimension to 5 product rating dimensions. The details are shown in Figure 7.4.

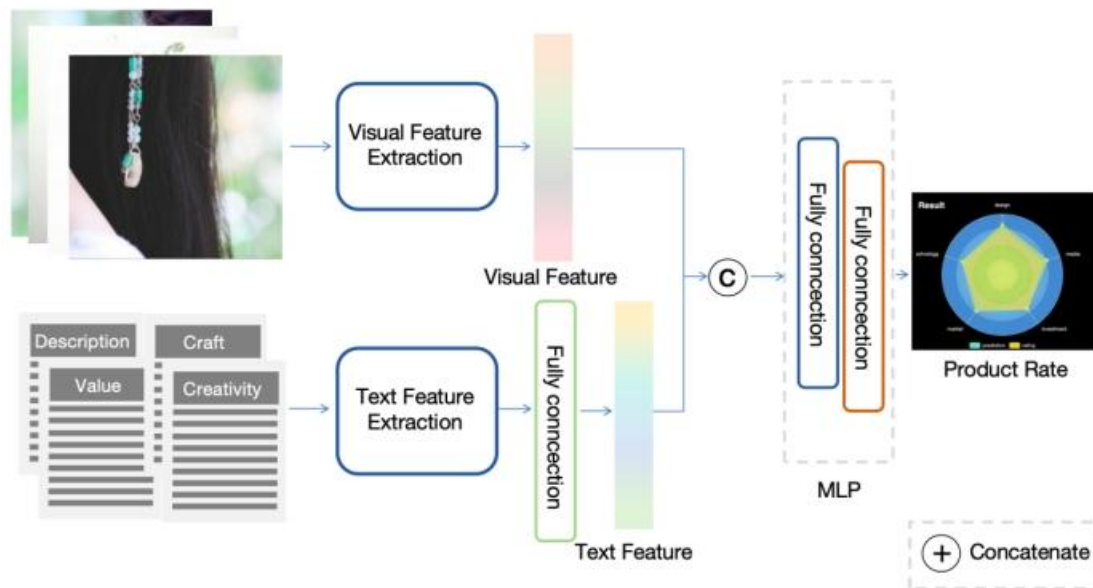


Figure 7.4 Vision-Language Fusion for Product Rating Algorithm

7.2.2 Training and Loss Function

The training session system is i5-12490f, Nvidia 3060Ti G6X, Asrock B660M-ITX, RAM 64G. At first, this Study used normal distribution with 0 for mean and 1.0 for standard deviation to initialize the parameters. Then, each model was trained in 16 epochs on the hardware system. The Adam with (0.9, 0.999) for betas was used during the training session, 0.001 for learning rate, without any weight decay and early stop.

Study 3 also used Mean Absolute Error (MAE) as the training criterion, as shown below. The y_{gt} is the five dimensions vector representing the value of media, investment, market, technology, and design.

$$l_n = \sum |y_{gt} - y_{pred}|, \#(7)$$

This y_{pred} is the corresponding prediction.

7.2.3 Performance Metric

In this work, the five famous metrics were be used to measure how well this evaluation model performance is, including ACC (Binary Classification Accuracy), SRCC (Spearman's rank-order correlation), PLCC (Pearson linear correlation coefficient), RMSE (Root mean square error) and MAE (Mean absolute error).

ACC is shown in the following formula. The y is the one-hone label. The $\hat{y}$ is model prediction. The TN is true negative. The TP is true positive. The FN is false negative. The FP is False positive.

$$ACC(\hat{y}, y) = \frac{TN + TP}{TN + FP + FN + TP}, \#(8)$$

SRCC is shown in the following formula. The $d_i = R(X_i) - R(Y_i)$ is the difference in rank between the label and prediction. The n is the number of samples.

$$SRCC(X, Y) = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \#(9)$$

PLCC is shown in the following formula. The x_i and y_i are prediction and label. The $\bar{x}$ and $\bar{y}$ are corresponding mean values. The s_x and s_y are the corresponding standard deviations, and the n is the number of samples.

$$PLCC(X,Y) = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y}, \# (10)$$

RMSE is shown in the following formula. The $\hat{y}$ and y are labels and predictions. Moreover, the N is the number of samples.

$$RMSE(\hat{y}, y) = \frac{1}{N} \sqrt{\sum_{N=1}^n (\hat{y} - y)^2}, \# (11)$$

MAE is shown in the following formula. The $\hat{y}$ and y are labels and predictions. Furthermore, the N is the number of samples.

$$MAE(\hat{y}, y) = \frac{1}{N} \sum_{N=1}^n |(\hat{y} - y)|, \# (12)$$

7.2.4 Dataset Quality Control and Deviation Analysis

This study implemented strict data quality control measures during the training process of the AI model. The research mainly focuses on three aspects: data preprocessing quality control, potential deviation analysis, and model development limitations.

The quality control of data preprocessing is mainly carried out through the following steps. Firstly, in the data cleaning stage, this study removes duplicate, invalid, and missing value samples, identifies and processes outliers. Then standardize the dataset to ensure consistency of data from

different sources. Finally, this study expands the training set through techniques such as rotation and scaling to enhance the data.

The training data may have the following biases. Firstly, there is a bias in sample selection, as the design projects mainly come from specific competition works. Next is the standard deviation of evaluation, as expert ratings are influenced by personal experience and are relatively subjective. Then there is data distribution bias, which means that some design types have smaller sample sizes. Finally, time deviation and the impact of design trends changing over time, as well as the fact that the dataset contains design works from several years ago, can also affect the evaluation results. To reduce the impact of data bias on model performance, this study adopted oversampling of minority class samples and achieved data augmentation by expanding the minority class samples. And invite multiple experts to cross evaluate to reduce subjective bias and achieve expert consensus.

Identify the following main limiting factors during the model development process. Firstly, there are limitations in computing resources. GPU memory capacity restricts batch processing size, training time is affected by hardware performance, and model complexity is constrained by computing power. Secondly, there are limitations in data quality, challenges in aligning multimodal data, and varying degrees of data standardization. Finally, and most importantly, there are limitations in the algorithm. The model has limited understanding of subjective evaluation indicators, and its cross domain generalization ability needs to be improved. The interpretability of the results also needs to be strengthened.

These limiting factors provide important references for future research directions. Subsequent research needs to focus on developing more efficient data preprocessing methods, exploring new model architectures to improve performance, and enhancing the interpretability and generalization ability of models.

7.3 Results

These studies emphasize the crucial role of deep learning, especially multimodal neural networks, in enhancing product conceptual design quality assessment. By utilizing advanced deep learning techniques, designers can achieve more comprehensive, objective, and effective evaluations, thereby developing innovative and aesthetically pleasing products that meet user needs and preferences.

7.3.1 Model Training Performance

This section will detail the performance of different models on various evaluation indicators. Moreover, it compares the advantages and disadvantages of single-modal and multimodal methods based on the training results of the model. CV models are Vgg16, ResNet18, Desene121, CustomizaNet, and ViT. BERT is an NLP model that takes input text as a description of the work, design concept, production process, and market value. Vgg16+Bert and ResNet18+Bert are multimodal models that combine images with text. This Study used these algorithms to train five evaluation metrics: Design, Technology, Market, Investment, and Media.

In the Design evaluation task, the multimodal models (ResNet18+BERT and VGG16+BERT) performed the best ACC, reaching 0.6311 and 0.618, respectively. The pure text model BERT (ACC=0.6382) outperforms all pure image models. This indicates that textual information may be more discriminative in design evaluation than visual features. However, all models performed weakly on this task, with SRCC and PLCC values low, which may reflect the subjectivity and complexity of design evaluation. The details are shown in Table 7.6.

Table 7.6 Model Training Results of Design

Method	ACC	SRCC	PLCC	RMSE	MAE
Vgg16	0.5023	-0.0168	0	0.1778	0.1478
ResNet18	0.5046	0.0473	0.0488	0.1792	0.1481
Desene121	0.5027	0.0181	0.025	0.1931	0.1507
CustomNet	0.5609	0.1985	0.2032	0.1756	0.1447
ViT	0.5446	0.1482	0.1601	0.1821	0.1502
Bert	0.6382	0.2907	0.3029	0.1892	0.1499
Vgg16+Bert	0.618	0.3196	0.3345	0.1856	0.1488
ResNet18+Bert	0.6311	0.3304	0.3448	0.1749	0.1382

In the Technical evaluation task, the ResNet18+BERT model performed the best (ACC=0.7301, SRCC=0.4889, PLCC=0.4587) 。 It is worth noting that pure image models such as VGG16 and Custom Net perform well on ACC but have lower SRCC and PLCC values, which may suggest room for improvement in sorting and linear correlation of the model. Overall, the performance of technical evaluation tasks is better than that of design evaluation, which may be due to the objective and quantifiable technical characteristics, as shown in Table 7.7.

Table 7.7 Model Training Results of Technology

Method	ACC	SRCC	PLCC	RMSE	MAE
Vgg16	0.7378	-0.0041	0	0.1788	0.1424
ResNet18	0.7359	0.3467	0.3186	0.169	0.1313
Desene121	0.7368	0.2361	0.1776	0.1758	0.1394
CustomNet	0.738	0.367	0.3412	0.1677	0.1288
ViT	0.6651	0.2427	0.2247	0.1806	0.1452
Bert	0.7178	0.44	0.407	0.1722	0.1295
Vgg16+Bert	0.7155	0.4591	0.4202	0.1685	0.1273
ResNet18+Bert	0.7301	0.4889	0.4587	0.1635	0.1235

In the Market evaluation task, the ResNet18+BERT model performed the best again (ACC=0.723, SRCC=0.5874, PLCC=0.5718), as shown in Table 7.8. The pure text model BERT performs outstandingly in this task, indicating that textual descriptions may contain key information in market evaluation. The advantages of multimodal models are evident and effective in this task, possibly because market evaluation requires a comprehensive consideration of the visual and descriptive features of the product.

Table 7.8 Model Training Results of Market

Method	ACC	SRCC	PLCC	RMSE	MAE
Vgg16	0.5825	-0.0214	0	0.1814	0.1535
ResNet18	0.6236	0.3757	0.3705	0.1707	0.1415
Desene121	0.5079	0.2619	0.1841	0.1928	0.1583
CustomNet	0.6549	0.4223	0.4189	0.1639	0.1343
ViT	0.598	0.3121	0.314	0.1802	0.1482
Bert	0.6999	0.5385	0.5219	0.1612	0.1255
Vgg16+Bert	0.7195	0.5668	0.5479	0.1598	0.1216
ResNet18+Bert	0.723	0.5874	0.5718	0.1523	0.117

In the Investment evaluation task, the ResNet18+BERT model still performed the best (ACC=0.7251, SRCC=0.4897, PLCC=0.4805). However, all models' RMSE and MAE values are high, which may reflect the high uncertainty and complexity of investment evaluation tasks. The advantages of text-only and multimodal models are pronounced in this task because investment decisions usually require more contextual and descriptive information. The details are shown in Table 7.9.

Table 7.9 Model Training Results of Investment

Method	ACC	SRCC	PLCC	RMSE	MAE
Vgg16	0.6807	0.0218	0	0.2094	0.1807
ResNet18	0.6797	0.2592	0.2558	0.2022	0.1712
Desene121	0.6592	0.1856	0.161	0.2093	0.1813
CustomNet	0.6688	0.2743	0.2741	0.201	0.1696
ViT	0.5859	0.1867	0.1838	0.2178	0.1852
Bert	0.7036	0.4181	0.4003	0.2069	0.1574
Vgg16+Bert	0.7122	0.4678	0.4589	0.198	0.1522
ResNet18+Bert	0.7251	0.4897	0.4805	0.1901	0.1467

In the Media evaluation task, the ResNet18+BERT model achieved the highest ACC (0.8051) and PLCC (0.4046), as shown in Table 7.10. Interestingly, pure image models such as VGG16 and Desene121 perform well on ACC but have extremely low SRCC and PLCC values, suggesting challenges in capturing subtle differences in Media evaluation. Overall, the ACC value of media evaluation tasks is the highest, but the correlation index is low, which may reflect the special nature of Media evaluation.

Table 7.10 Model Training Results of Media

Method	ACC	SRCC	PLCC	RMSE	MAE
Vgg16	0.7841	0.0094	0	0.1646	0.1329

ResNet18	0.7762	0.1341	0.1373	0.1686	0.1395
Desene121	0.7845	0.0446	0.0231	0.1677	0.128
CustomNet	0.7843	0.1961	0.2133	0.1598	0.128
ViT	0.7574	0.1647	0.1894	0.1714	0.1373
Bert	0.796	0.3303	0.3717	0.1665	0.1281
Vgg16+Bert	0.7889	0.3422	0.3782	0.1669	0.1272
ResNet18+Bert	0.8051	0.3549	0.4046	0.157	0.1202

This study explores the performance of multimodal deep learning models in product design evaluation, combining image and text information to provide a more comprehensive and accurate assessment. The experimental results show that the multimodal model performs the best in most evaluation tasks, especially achieving high accuracy in predicting the overall and attribute level ratings of products.

7.3.2 Result Analysis

This Study explores the performance of different artificial intelligence models in product design evaluation tasks, covering five evaluation dimensions: Design, Technology, Market, Investment, and Media. The experimental results showed that the multimodal models, ResNet18+BERT and VGG16+BERT, performed the best in most evaluation tasks, indicating that combining visual and textual information can provide more comprehensive and accurate evaluation results.

The pure text model BERT outperforms the pure image model in most tasks, highlighting the importance of product description in the evaluation process.

This finding is consistent with previous research that suggests textual information may contain more key details related to evaluation.

The performance differences between different evaluation dimensions reflect the characteristics and challenges of each field. For example, the overall performance of Design evaluation tasks is weak, which may be due to their high subjectivity. In contrast, the performance of Technology and Media evaluation tasks is better because these fields have more objective evaluation criteria.

Figure 7.5 is an example showing the training results of the model. Due to copyright issues, the product images were intentionally blurred. The press is the algorithm's predicted score. GT is the true score of experts.

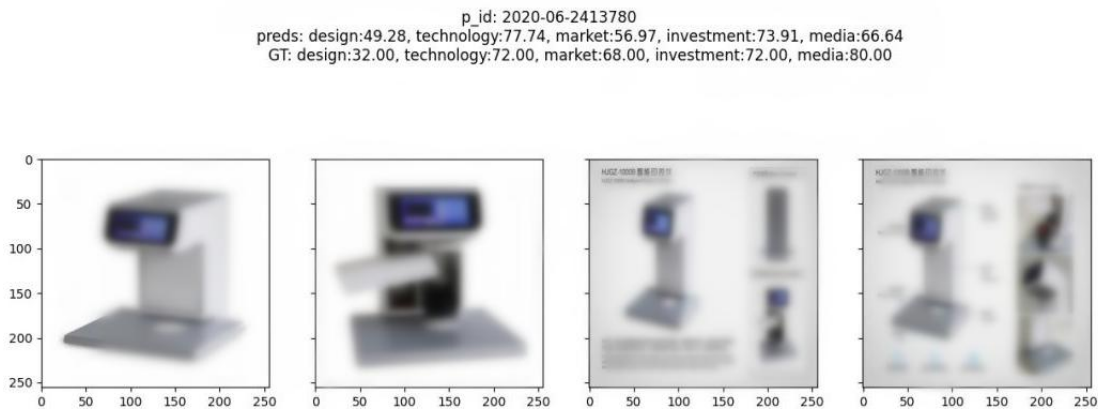


Figure 7.5 An Example of Training Results

It is worth noting that models perform well in accuracy (ACC) but poorly in correlation metrics (SRCC and PLCC), suggesting that we need to consider multiple evaluation metrics comprehensively to evaluate model performance. Future research can focus on improving the ability of models to capture subtle

differences and enhance correlation metrics, particularly in design and media evaluation tasks. In addition, considering the complexity of the evaluation process, methods that combine AI evaluation results with human expert judgment can be explored in the future to improve the reliability and acceptance of the evaluation.

Overall, this study provides important insights into the potential and limitations of AI applications in product design evaluation, laying the foundation for future research and applications in this field.

7.4 Summary and Discussions

In the field of design evaluation, integrating multimodal datasets and deep learning techniques has become a key focus for researchers to improve the efficiency and accuracy of the design process. Multimodal datasets combine data types such as text, images, and gestures to provide a more comprehensive understanding of design elements (Zhou et al., 2018). These datasets are particularly valuable in simulating real-world scenarios, such as international online shopping, where multiple patterns are used to describe products (Zhou et al., 2018). However, using multimodal data may lead to increased complexity, inference time, and memory requirements of the model, so it is necessary to thoroughly examine the advantages and differences compared to single modal models (Thißen & Hergenröther, 2023). Deep learning plays a crucial role in processing multimodal data for design evaluation, such as convolutional neural networks (CNN) and recurrent neural networks (RNN) (Lu et al., 2022). For example, CNN excels at automatically extracting complex features from multimodal gesture data, enabling more detailed analysis and interpretation (Lu et al., 2022). In

addition, incorporating artificial intelligence (AI) into the design process has fundamentally changed the generation of conceptual designs through deep learning algorithms (As et al., 2018). These algorithms can extract necessary design components based on performance standards and recombine them to generate novel design solutions, demonstrating the transformative potential of artificial intelligence in design innovation (As et al., 2018).

The application of artificial intelligence in design evaluation is not limited to physical products but extends to software products in fields such as medicine and education. In the medical field, artificial intelligence diagnostic tools driven by deep learning algorithms have demonstrated outstanding capabilities in automatically classifying medical conditions from imaging data (Lakhani & Sundaram, 2017). By enhancing the dataset and utilizing convolutional neural networks, researchers have achieved significant improvements in the accuracy of the diagnostic process, highlighting the potential of artificial intelligence to revolutionize medical imaging analysis (Lakhani & Sundaram, 2017). Similarly, in the educational environment, artificial intelligence technology has reshaped the paradigm of instructional design, leading to a shift in teaching objectives, content delivery, and evaluation methods (Jian et al., 2021). The integration of artificial intelligence has ushered in a new era of personalized and interactive learning experiences, meeting diverse student needs, and improving educational outcomes. In addition, integrating artificial intelligence and multimodal datasets has paved the way for innovative methods in business model design and transformation management for entrepreneurial enterprises (Fang, 2023). By leveraging insights driven by artificial intelligence, entrepreneurs can design agile business models, capitalize on emerging opportunities, and, more precisely, navigate complex market environments (Fang, 2023). Artificial intelligence

algorithms have played a significant role in optimizing coordination delivery routes, reducing indirect costs, and improving operational efficiency in supply chain management (Yao, 2020). Through adaptive weighting methods and route optimization techniques, artificial intelligence enables organizations to simplify delivery processes and improve overall coordination performance (Yao, 2020).

In summary, the synergistic effect between multimodal datasets, deep learning applications, and artificial intelligence technologies has redefined the design evaluation landscape in various fields. From architecture to medicine, education, and entrepreneurship, integrating AI-driven solutions has opened new possibilities for creativity, efficiency, and innovation. By utilizing the power of AI to analyze various data modalities, designers and decision-makers can make wise choices and drive transformative design. Shaping and developing intelligent design evaluation models is at the forefront of future research. Based on relevant research, this Study explored the performance of various artificial intelligence models in product design evaluation tasks. The experimental results of Study 3 are summarized as follows.

- The multimodal fusion model performs the best. In all evaluation dimensions (design, technology, market, investment, media), multimodal models that combine images and text, such as ResNet18 with BERT, outperform single-modal models. This indicates that considering the product's visual and textual descriptive information can provide more comprehensive and accurate evaluation results.

- NLP algorithm is superior to CV models. The BERT model performs better than pure image models in most evaluation dimensions. This indicates that the textual description of the product, such as design creativity, production process, market value, etc., may contain more key information related to evaluation.
- There are performance differences between evaluation dimensions. There are significant differences in the performance of models across different evaluation dimensions. For example, in the "design" dimension, the performance of all models is low (ACC highest at 0.6311), while in the "media" dimension, the models perform well (ACC highest at 0.8051). This reflects the characteristics and difficulty differences of different evaluation dimensions.
- There are challenges in the artificial intelligence evaluation of design indicators. All models perform poorly in the design dimension. This may reflect the subjectivity and complexity of design evaluation, implying the challenges that AI models face when dealing with highly subjective evaluation tasks.

These findings provide important insights into the potential and limitations of AI applications in product design evaluation. However, current methods also have shortcomings. Firstly, there is still room for improvement in the model's performance in the design dimension due to the strong subjectivity of design evaluation. Secondly, the experiment lacks an interpretability analysis of the model decision-making process, making it difficult for users to understand the scoring criteria. The correlation between different evaluation dimensions

was not considered, and each dimension was trained independently. Finally, the dataset size may need to be expanded to improve the model's generalization ability.

In response to these shortcomings, future research directions can be proposed. Considering the complexity of the evaluation process, the next Study needs to explore methods that combine AI evaluation results with human expert judgment to improve the reliability and acceptance of the evaluation. Subsequent research requires conducting human-machine comparative experiments to evaluate the performance differences between AI models and human experts in designing evaluation tasks. Study 4 will use empirical methods to effectively integrate AI evaluation tools into actual design workflows to improve designers' practicality and acceptance.

This chapter introduces designing and implementing an AI evaluation model based on multimodal datasets. By developing an AI algorithm that combines visual feature extraction, text analysis, and multimodal fusion, this Study provides an innovative approach to product design evaluation. Although the model may have specific limitations, its development provides a technical basis for the next human-computer comparison experiments (Study 4), enabling researchers to verify the AI model's evaluation performance further. The next chapter will verify the effectiveness of this model through human-computer comparison experiments.

Chapter 8. Study 4- Human-machine Comparative Evaluation Experiment

This chapter describes the design and implementation of a human-machine comparative evaluation experiment. Based on the AI model developed in Study 3, this study designed a website interface, the Artificial Intelligence Assisted Product Design Evaluation System (AiDE), that allows human experts and AI models to evaluate product designs in parallel. Study 3 showed that the method combining image and text information performed the best in the evaluation task. This experiment aims to verify the evaluation performance of the AI model and explore the possibility of human-machine collaborative evaluation. The core content is to evaluate the accuracy and reliability of AI models through human-machine comparison experiments and compare the performance differences between AI and human experts in evaluation tasks. It helps understand the strengths and weaknesses of AI models and offers important insights for improving AI-assisted design evaluation tools. The experimental results verify the AI model's effectiveness and provide important input for the prototype verification study (Study 5), which is crucial for the practical application of AI tools.

8.1 Introduction

It is crucial to consider a range of factors that may affect the performance and results of artificial intelligence (AI) product design evaluation models to evaluate their effectiveness. Al-Sa'Di and Miller emphasize that a key factor is integrating AI tools into the design process. Their research explores the impact of artificial intelligence language models (such as ChatGPT) on user experience in the design thinking process, emphasizing how this integration

enhances the design decision-making process and has the potential to improve the overall user experience of products (Al-Sa'di & Miller, 2023).

In addition, a study on product design decision models based on deep residual networks emphasizes the importance of developing AI design decision models to improve efficiency and reduce the impact of decision preferences on product design and development (Zong & Wang, 2022). By utilizing advanced AI models such as deep residual networks, evaluation models can benefit from more accurate and data-driven decision-making processes, thereby enhancing design outcomes. As Zhang et al. (2022) discussed, consumer response to artificial intelligence design is crucial for evaluating the effectiveness of AI-based product design. Understanding consumers' willingness to pay for products designed with artificial intelligence can provide insights into market acceptance and perceived product value. These pieces of information are of great significance for evaluating the success of evaluation models in meeting consumer expectations and preferences.

Moreover, the study by Liu (2015) on collaborative modeling and design of intelligent product networks emphasizes how artificial intelligence algorithms can effectively promote cognitive thinking storage and feedback mechanisms in product design optimization. This study utilizes artificial intelligence algorithms for cognitive tasks and artistic image optimization, and the evaluation model can potentially enhance the creativity and aesthetic appeal of design products, thereby improving overall design quality. Yang et al. (2021) researched intelligent product form design based on cognitive dynamics and spider web structure, introducing the concept of selecting feasible solutions from intelligent design sketches. This method is based on

the designer's cognition, and integrating human-centered design principles and cognitive processes into AI-driven design systems helps to improve evaluation models, resulting in more user-centered and innovative product designs.

Oh et al. (2019) proposed a framework for deep generative design. This framework integrates topology optimization and generative models to generate aesthetically pleasing and engineering-optimized design options. This study utilizes advanced generative design techniques to evaluate models that can provide a range of visually appealing and functionally optimized design alternatives, thereby enhancing the overall design evaluation process. Verganti et al. (2020) summarized how artificial intelligence can create user-centric solutions that are constantly updated through learning iterations. Artificial intelligence's continuous learning and adaptability can ensure that design decisions are guided by real-time data and user feedback, significantly affecting the effectiveness of evaluation models, and resulting in more personalized and innovative product designs.

In summary, evaluating the effectiveness of AI product design evaluation models requires a thorough analysis of factors such as integration with AI tools and cognitive design considerations. By drawing insights from these different perspectives, designers, and researchers can evaluate the impact of artificial intelligence on product design results and improve evaluation models to enhance their effectiveness in creating innovation and user centricity.

In the early stage of this study, a literature review was conducted to identify research gaps in AI-assisted product design evaluation, and qualitative and quantitative methods were used to explore designers' needs and expectations for AI tools. Subsequently, we constructed a multimodal product design dataset and designed and implemented an AI evaluation model based on deep learning. The model evaluation results indicate that the multimodal method combining image and text information performs the best in product design evaluation tasks.

To further validate the practicality of the AI evaluation model and promote human-machine collaboration, this chapter will introduce a web-based graphical user interface (GUI) design. This interface aims to provide designers with an intuitive and easy-to-use platform to use AI models for product design evaluation easily. The design of GUI considers user experience and interaction design principles to ensure that it conforms to the designer's workflow and habits.

The core content of this chapter is to evaluate the accuracy and reliability of AI models through human-machine comparative experiments. This experiment compares the performance differences between AI models and human experts in product design evaluation tasks. By having experts and AI models rate a series of product samples, this study can evaluate the accuracy and consistency of the model. This comparison helps researchers understand the advantages and limitations of AI models and provides valuable insights for improving AI-assisted design evaluation tools in the future. This study also verifies the application prospects of AI in product design evaluation through empirical methods. The research findings also provide an important

basis for constructing a more comprehensive and effective theoretical framework for design evaluation.

8.2 Methods

This study adopts an innovative online experimental method to conduct human-machine comparative evaluation. This study developed and deployed a web platform (<http://118.195.135.213/ArtRate/dist>) called the Artificial Intelligence Design Evaluation System (AiDE) as a functional prototype of the experiment. The platform integrates multi-role user management, multimodal data processing, and human-machine collaborative evaluation.

The experiment invites junior designers to submit their product design projects, which are evaluated in parallel by an expert panel and an AI model. The evaluation dimensions include five aspects: design, technology, market, investment, and media. A sliding bar rating method of 0-100 points is used. This design not only ensures the comprehensiveness of the evaluation but also provides a basis for direct comparison of human and machine evaluation results. With this method, this study aims to evaluate the accuracy and reliability of the AI model rigorously. At the same time, Study 4 also explores the potential of AI in assisting or enhancing human evaluators, especially in terms of efficiency when dealing with large-scale data sets.

8.2.1 GUI Design and Website Function Implementation

Based on previous research results, study 4 has developed and launched a function prototype for empirical experiments, the Artificial Intelligence Design Evaluation System (AiDE).

8.2.1.1 GUI Design

The AiDE Artificial Intelligence Design Evaluation System is an innovative product design evaluation platform that integrates advanced concepts such as multi-role user management, multimodal data processing, and human-machine collaboration evaluation. If users are interested in the AiDE artificial intelligence design evaluation system, please log in to the website: <http://118.195.135.213/ArtRate/dist>. After registration, users should email jing86.luo@ to apply for relevant permissions, as shown in Figure 8.1. Researchers will grant relevant permissions. After granting expert permissions, users can upload design projects and rate them on the website.



Figure 8.1 Registration Page

One of the core design concepts of this system is multi-role user design, including four roles, ordinary users, expert users, staff users, and superusers, to meet the needs of different user groups. This layered design not only reflects the refined management of user permissions and functional access in the system but also helps to ensure the security of data and the fairness of the evaluation process, as shown in Figure 8.2.

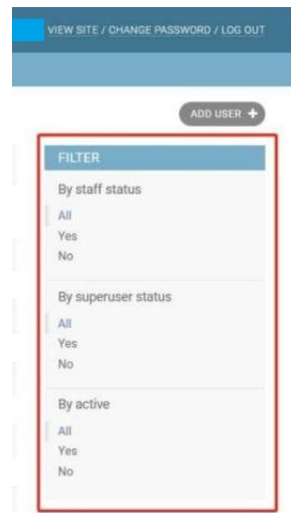


Figure 8.2 User Permission Setting Function

As shown in Figure 8.3, the AiDE system adopts a simple and clear design style, using a dark background with white text and icons to highlight key information and reduce visual interference effectively. The main functional modules, such as the homepage, user upload, and gallery, are clearly visible, making it easy for users to locate the required functions quickly. This design not only improves the usability of the system but also reflects the emphasis on user experience.

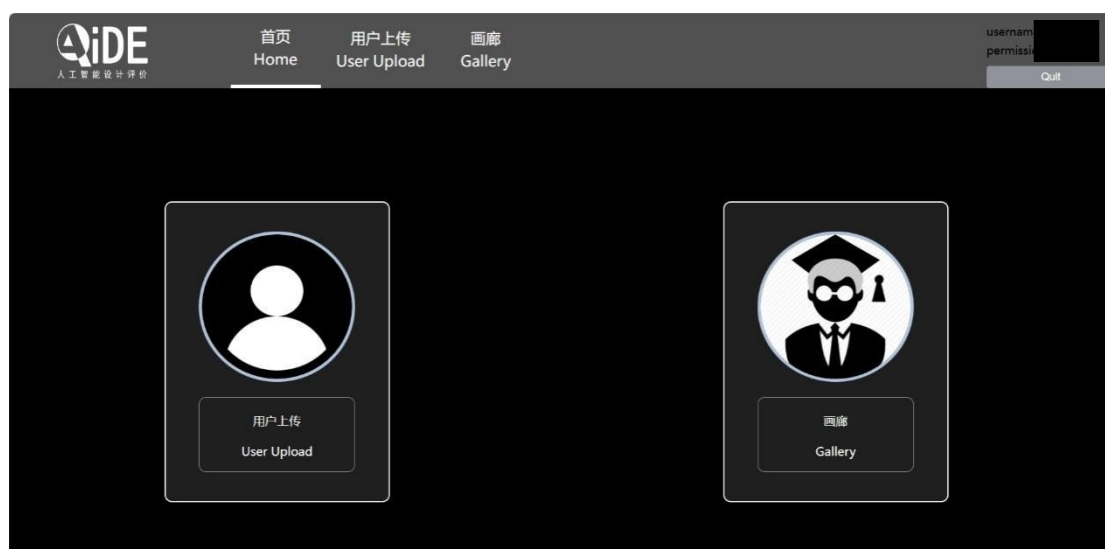


Figure 8.3 User Interface Design

Figure 8.4 shows the user upload interface, which supports image upload and text description input. This interactive form reflects the multimodal data input function of the system. This design helps to capture the visual and conceptual features of product design, providing more comprehensive data support for subsequent evaluations.

Figure 8.4 The User Upload Design Project Interface

The visualization display of evaluation results is another important feature of the AiDE system. As shown in Figure 8.5, the system utilizes data visualization techniques such as radar charts and bar charts to visually display evaluation results from multiple dimensions, including design, technology, market, investment, and media. Prediction (green) is the algorithm rating. Rating (yellow) is the rating given by human experts. This multidimensional visualization display helps users quickly understand the evaluation results and facilitates multidimensional comparisons, providing dedicated support for design decisions.



Figure 8.5 Visualization of Rating Results

The system is designed with two modes, expert rating, and machine rating, reflecting the evaluation concept of human-machine collaboration. Figure 8.6 shows the expert rating interface, where experts can rate products based on dimensions. This design utilizes the efficiency and consistency of AI and retains the professional judgment of human experts, which is expected to improve the accuracy and credibility of evaluations.

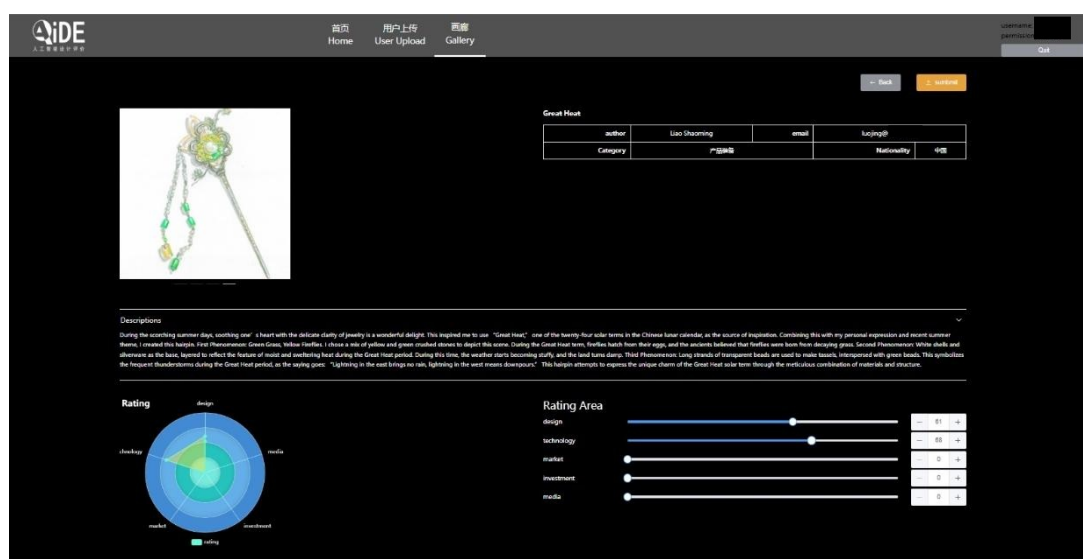


Figure 8.6 Expert rating process

Figure 8.7 shows the gallery interface of the AiDE system. The gallery highlights multiple product design works; each presented as images with titles and author information. The search bar at the top of the interface allows users to search for specific works by keywords or dates, improving the usability and efficiency of the system.

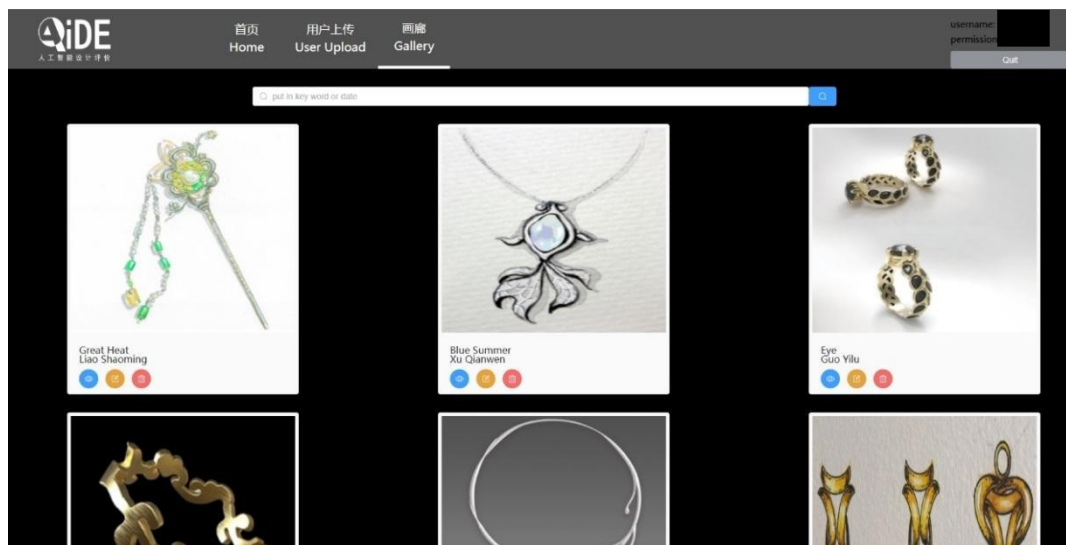


Figure 8.7 Gallery

The system uses clear icons and labels to represent distinct functions like viewing details, expert ratings, and machine ratings. This design helps users understand the scope of AI system capabilities and conforms to the design principles of helping users understand AI functions. The system provides options for rating and deleting products, emphasizing collecting, organizing, and utilizing user feedback, an essential aspect of AI system design.

As shown in Figure 8.8, the system clearly labels the machine rating options, allowing users to know when to use AI for evaluation. This transparency

helps establish user trust and complies with the design principles of explainable artificial intelligence.

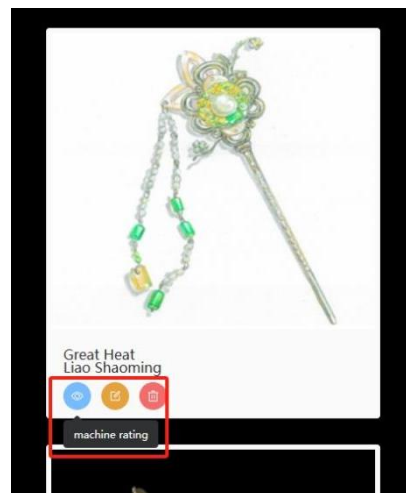


Figure 8.8 Machine Rating icon

Figure 8.9 represents the expert rating function. The interface provides two options: expert rating and machine rating, reflecting the evaluation method that combines human professional knowledge with AI capabilities. This design reflects the principle of integrating human expertise and AI capabilities in decision-making. By providing different rating options and clear functional labels, the system strives to improve its interpretability, which helps establish user trust in the AI system.

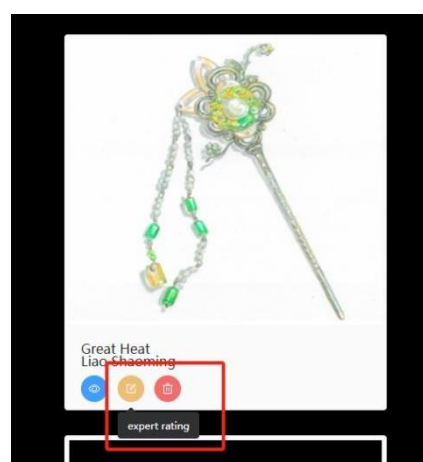


Figure 8.9 Expert Rating icon

As shown in Figure 8.10, the interface provides the option to delete products. The icon features a red and trash can pattern, with a reminder artwork. This indicates that users have a certain degree of control over the system content. This is an important aspect of user-centered AI design.

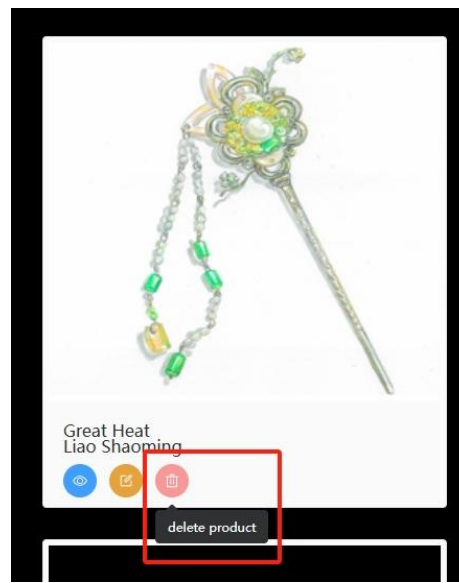


Figure 8.10 Delete Product icon

This multimodal data display method helps to comprehensively capture the product design's visual and conceptual features, providing more comprehensive data support for subsequent evaluations. Users can enter the detailed evaluation page by clicking on the work, reflecting the hierarchical design of the system.

In terms of system management, as shown in Figure 8.11 and Figure 8.12, AiDE has designed a backend management system based on Django. This design helps researchers to manipulate and manage data such as product design works and ratings. This design also facilitates data maintenance and

system management for administrators, providing convenience for collecting and analyzing research data.



Figure 8.11 Backend Login

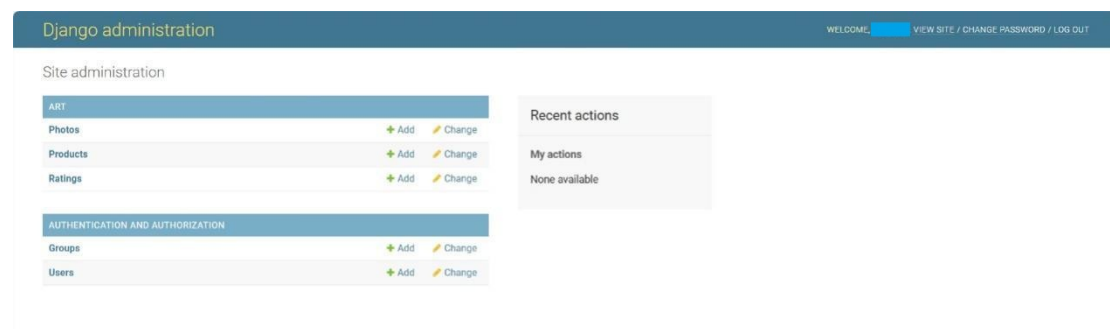


Figure 8.12 Backend Management

Overall, the design philosophy and interface layout of the AiDE system emphasize simplicity, visualization, and user-friendliness. This is all based on previous research, Study 1- Study 4. Its interaction design simultaneously considers advanced concepts such as multi-role requirements, multimodal data processing, and human-machine collaboration evaluation. This design is suitable for practical applications in product design evaluation and provides powerful tool support for related academic research. By integrating artificial intelligence technology with traditional design evaluation methods, the AiDE system provides new possibilities for innovation and development in product design. It is expected to promote the progress of design evaluation methods and design quality improvement.

8.2.1.2 Website Function Implementation

AiDE is a web platform for product data collection and expert and artificial intelligence scoring. The system architecture is shown in Figure 8.13, adopting a modern front-end and backend separation design pattern. This combination can provide an efficient user interface and powerful data processing capabilities. The front-end framework uses Vue.js, a progressive JavaScript framework for building user interfaces. The front-end design of this platform includes four main modules: Machine Rating, Expert Review, Gallery, and User Upload. This modular design reflects the system's pursuit of user-friendliness and functional clarity.

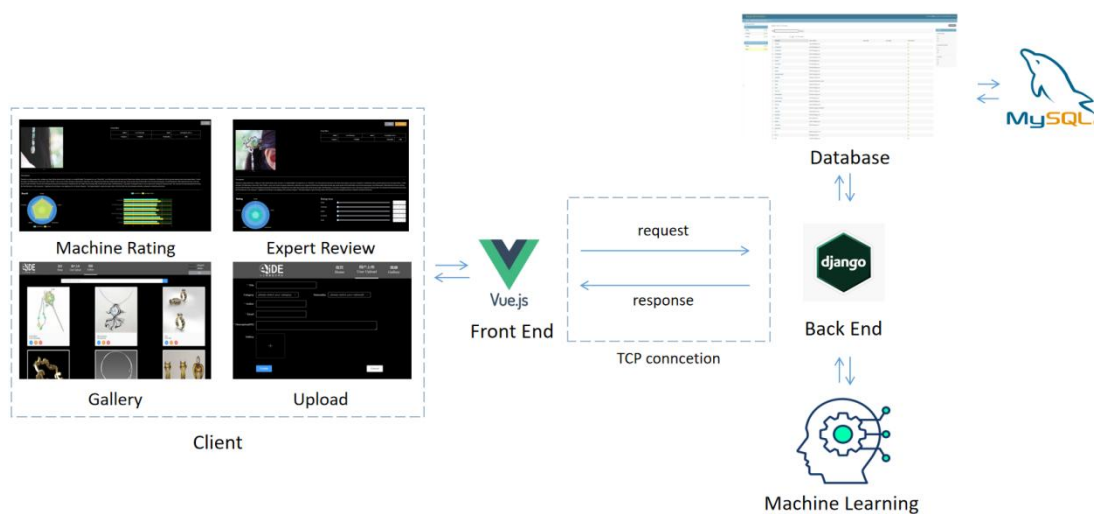


Figure 8.13 Website Framework

During the User Upload process, the system transmits information as form data. This method can effectively encapsulate complex product information, including text and image data. The Django backend achieves efficient data storage and processing by parsing frame structures while generating thumbnails, which helps optimize system performance and user experience.

The design of the Gallery module adopts a pagination display method. This method can effectively manage a large amount of product data and improve page loading speed and user experience. The backend returns information for nine products at once. This batch-processing method can reduce server load and improve system response speed. The Expert Review module embodies the evaluation concept of human-machine collaboration. The system provides comprehensive product information, allowing expert users to evaluate products from multiple dimensions (design, technology, market, investment, and media). This multidimensional evaluation method is consistent with the current research trend in product design and evaluation.

The Machine Rating module demonstrates the application of artificial intelligence in product evaluation. The system can automatically rate multiple dimensions of the product, which not only improves evaluation efficiency but also provides users with objective reference standards. Visualizing the results helps users better understand and compare the evaluation results. The server-side implementation utilizes Django, an advanced web development framework based on Python known for its ability to facilitate rapid deployment of secure and scalable web applications. The system is deployed on Tencent Cloud servers configured with four cores and 16GB of memory, which can meet the needs of medium-sized web applications.

In terms of data storage, AiDE uses MySQL relational database and adopts a non-distributed storage method considering the current data volume. The front-end and back-end data exchange is achieved through the HTTP protocol, the most used data transmission protocol for web applications. This architecture design reflects user experience, system performance, and scalability considerations. The system can provide a more flexible user

interface and efficient data processing capabilities by adopting a front-end and back-end separation architecture. Meanwhile, choosing mainstream technology stacks, Vue.js, Django, and MySQL, is also beneficial for system maintenance and future expansion. The design and development of the AiDE platform fully utilizes modern web technology and artificial intelligence algorithms, providing a comprehensive solution for product data collection and evaluation. This method, which combines manual expert evaluation and machine automatic scoring, has the potential to bring new research directions and practical applications to the field of product design and evaluation. In addition, the design philosophy of the AiDE platform reflects the application of human-machine collaboration in design evaluation. By combining machine scoring and expert review, the system has the potential to provide more comprehensive and objective product evaluation results, which is in line with the current research trend of artificial intelligence-assisted design.

Overall, the design and architecture of the AiDE platform reflect adherence to best practices in modern web application development, as well as innovative attempts to integrate artificial intelligence technology in the field of design evaluation.

8.2.2 Human-machine Comparison Experiment Design

This study adopts a mixed method design, combining quantitative and qualitative data collection techniques, aiming to evaluate the accuracy of the AI-assisted product design evaluation system (AiDE) evaluation results. The experimental design includes participant selection, experimental process, and human-machine comparison rating.

8.2.2.1 Participant Selection and Recruitment

The experimental participants were divided into two main groups: the product design group and the evaluation group. The product design team comprises junior students majoring in product design, and this selection criterion ensures that participants have sufficient professional knowledge to complete the task. The evaluation team comprises design experts with over four years of work experience. The recruitment process for experts utilized social media groups and professional networks. This stratified sampling and recruitment strategy helps ensure the typical diversity of the sample.

8.2.2.2 Experimental Process

The experimental process consists of two main stages: the product design stage and the evaluation stage. Figure 8.14 shows the overall experimental process of this study.

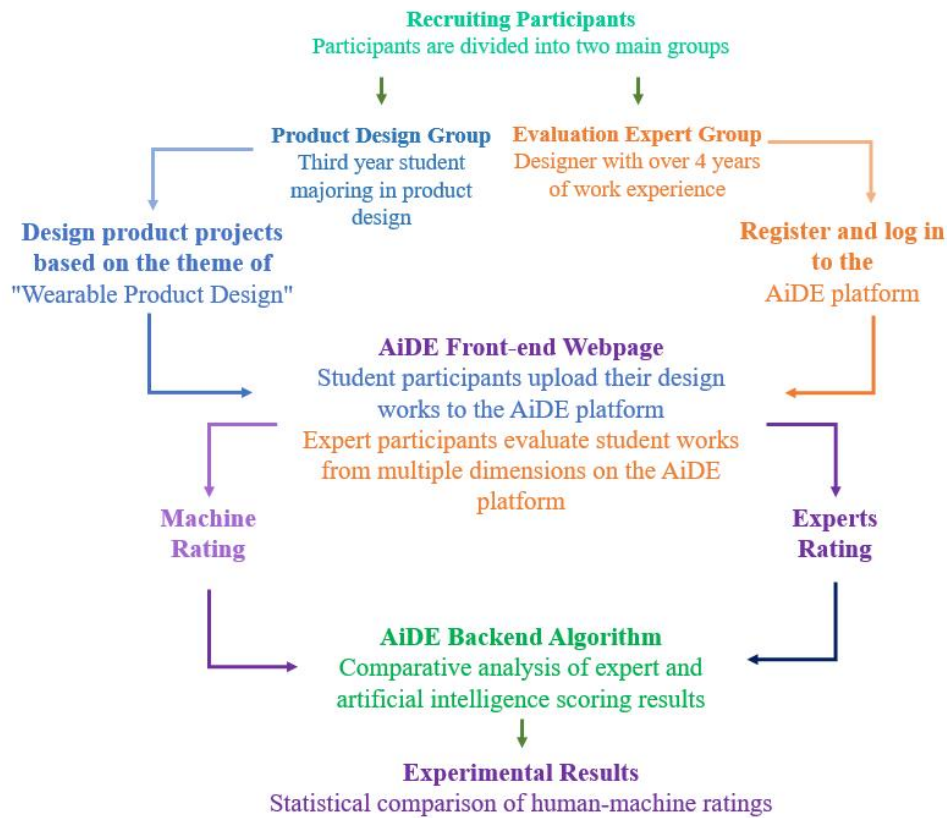


Figure 8.14 Experimental Process

In the product design stage, student participants are required to create design works based on wearable product design as the theme. Then, upload the design project, including renderings and specifications, to the AiDE platform.

During the evaluation stage, expert participants log in to the AiDE platform to rate student works from multiple dimensions, as shown in Figure 8.15. The scoring dimensions include five key aspects: design, technology, market, investment, and media. The expert rating adopts a sliding bar format with a range of 0-100 points, which allows evaluators to give precise scores.

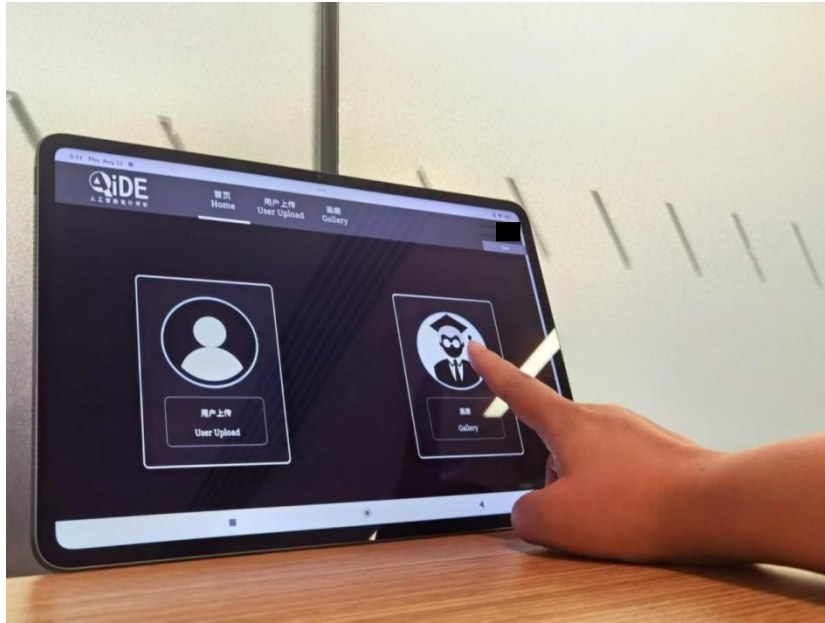


Figure 8.15 Experimental Procedures

8.2.2.3 Human-machine Comparison Rating

The AI evaluation algorithm of the AiDE platform adopts multimodal data analysis technology to process the effect drawings and design specifications uploaded by students. This method allows the system to capture product features comprehensively. AI models may also incorporate expert rating data for training and optimization, which reflects the current research trend of AI-assisted design.

The rating results are visualized through radar charts, which intuitively present the product's performance in various dimensions. The interface also displays rating records, including the evaluators' usernames and specific scores for each dimension, increasing the transparency and traceability of the evaluation process. It is worth noting that the AiDE platform supports manual evaluation and integrates artificial intelligence-assisted evaluation

functions. This design reflects the widespread application trend of artificial intelligence technology in intelligent manufacturing.

8.2.3 Evaluation Indicators

This study will use the evaluation metrics used in Section 6.2.3, Performance Metrics from Chapter Six, to compare and analyze human and artificial intelligence ratings. These indicators include the following.

- Binary classification accuracy (ACC): This indicator evaluates the model's overall performance in classification tasks, reflecting the proportion of correctly classified samples.
- Spearman rank correlation coefficient (SRCC): a non-parametric measure used to evaluate the ranking correlation between two variables, particularly suitable for assessing the ranking consistency between model predictions and human scores.
- Pearson Linear Correlation Coefficient (PLCC): This indicator measures the degree of linear correlation between model predictions and human scores, reflecting the overall trend consistency.
- Root Mean Square Error (RMSE): This indicator reflects the average deviation between model predictions and actual values and is more sensitive to larger errors.

- Mean Absolute Error (MAE): This indicator measures the average absolute difference between the predicted value and the actual value of the model and is less sensitive to outliers.

The comprehensive use of these indicators will enable a comprehensive evaluation of the performance differences between artificial intelligence models and human experts in product design evaluation tasks. This method not only checks classification accuracy but also focuses on ranking consistency and numerical accuracy of scores. Through these multidimensional comparisons, Study 4 can better understand the advantages and limitations of artificial intelligence models in simulating human evaluation processes, providing key evidence for further improving AI-assisted design evaluation tools.

8.3 Results

This study adopts a mixed method design, combining quantitative and qualitative data collection techniques, to evaluate the effectiveness of an AI-assisted product design evaluation system.

8.3.1 Website Usage Data Analysis

Based on the usage data and experimental statistics of the AiDE platform, this study conducts an in-depth analysis of participants' behavior patterns and user experience.

8.3.1.1 Participants and Sample

The AiDE platform recruited 129 users to participate in the experiment, including students from the product design group and experts from the evaluation group. The platform collected 61 design works, and 68 designers gave 1075 ratings, as shown in Figure 8.16 and Figure 8.17.



Figure 8.16 Number of Design Projects

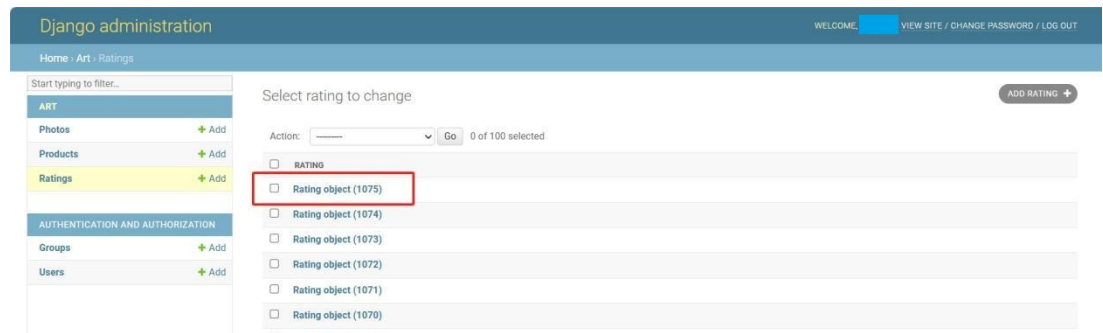


Figure 8.17 Number of Ratings

Each project has a maximum of 48 ratings, a minimum of 15 ratings, and an average of 20 ratings. This multi-person review method can more comprehensively capture various aspects of creativity, improving the reliability of the rating. Figure 8.18 shows one of the design projects as a rating sample.

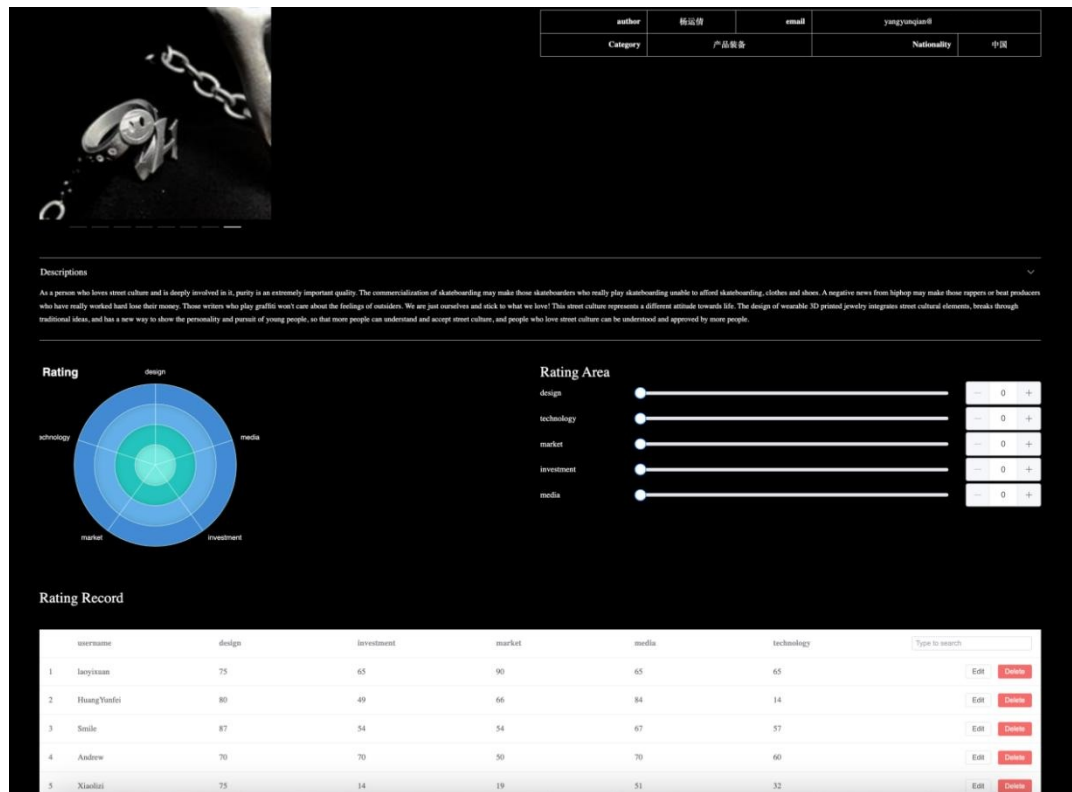


Figure 8.18 Rating Sample

8.3.1.2 Analysis of Exports Rating Behavior

According to experimental record data, online reviewers rate design projects at 3-5 per minute. The rating data of the AiDE platform (1075 ratings from 68 designers) reflects similar high efficiency. This rapid scoring method helps to filter out the most promising ideas.

The system has designed different user roles, including regular users, expert users, and administrators. This layered design reflects refined management of user permissions and functional access, which helps ensure data security and fairness in the evaluation process. The evaluators of the AiDE platform are design experts with work experience, which may improve the professionalism and reliability of the scoring.

The rating interface adopts a sliding bar form, with a range of 0-100 points, which allows evaluators to score accurately. The scoring dimensions include five aspects: design, technology, market, investment, and media, reflecting a comprehensive product evaluation. This multidimensional evaluation method has certain similarities with other dimensions used in research, such as creativity, clarity, novelty, practicality, and product value.

8.3.1.3 User Interface and Interaction Design

The interface design of the AiDE platform is simple and clear, with a dark background and white text, effectively reducing visual interference and improving user experience. The main functional modules (such as homepage, user upload, and gallery) are clearly visible, making it easy for users to locate the required functions quickly. The gallery interface adopts a grid layout to display design works, with each work presented in thumbnail form and accompanied by title and author information. This design facilitates users to browse and compare multiple works quickly.

8.3.1.4 Human-machine Collaboration Evaluation

The system supports manual evaluation and integrates the AI evaluation function. This design reflects the widespread application trend of artificial intelligence technology in intelligent manufacturing and is expected to improve the efficiency and objectivity of evaluation. Multimodal data input (image upload and text description) helps comprehensively capture product design's visual and conceptual features.

The comparison results between machine scoring and manual scoring can be divided into three categories: similar human-machine rating results, average performance of human-machine rating results, and significant differences in human-machine rating results. As shown in Figure 8.19, the machine rating is highly consistent with the expert rating in a sample with reliable performance. Prediction (green) is the algorithm rating. Rating (yellow) is the rating given by human experts. The results indicate that, in respects, artificial intelligence models have been able to simulate human experts' evaluation criteria accurately. This high consistency may occur in more objective evaluation dimensions, such as technical or functional indicators.



Figure 8.19 A Sample of Similar Human-machine Rating Results

The sample with average performance (Figure 8.20) shows a certain difference between machine and manual scoring, but the overall trend is similar. This situation reflects that there is still room for improvement in certain evaluation dimensions of artificial intelligence models, which may require more training data or complex algorithms to improve accuracy.

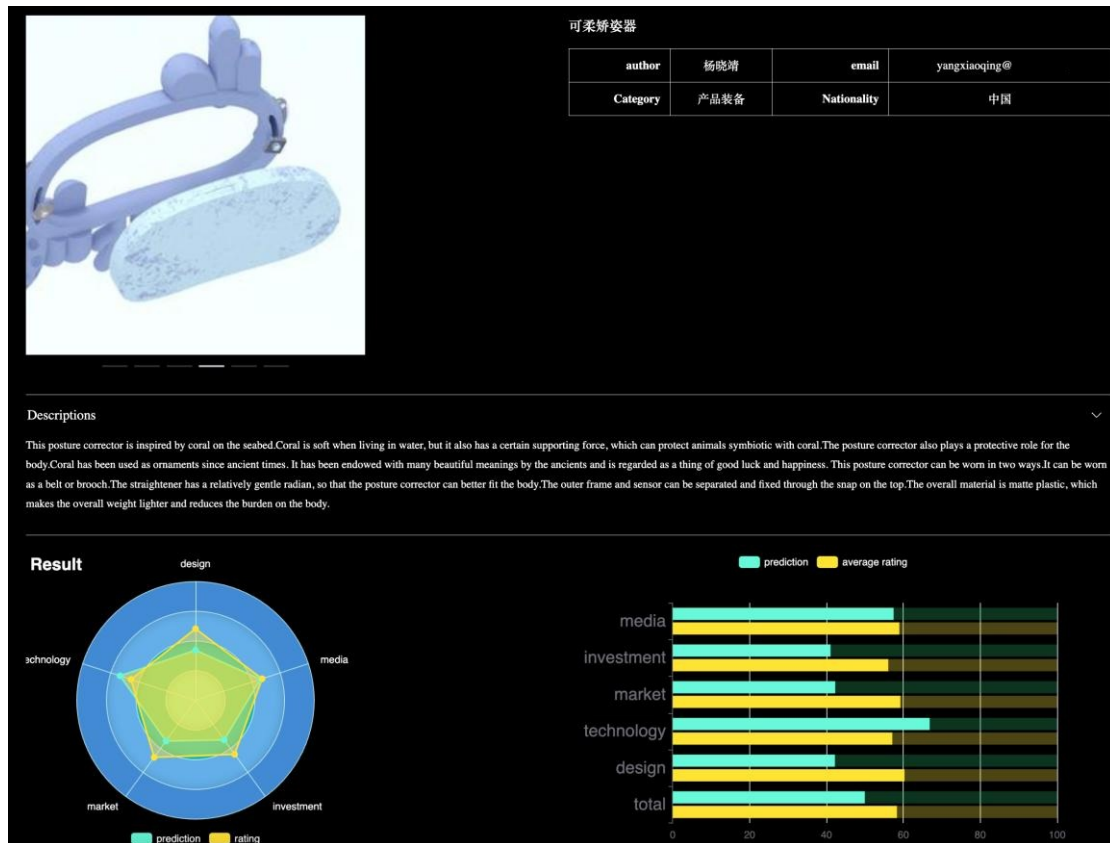


Figure 8.20 A Sample of Average Performance

The cases with deficient performance reveal a significant difference between machine scoring and manual scoring, as shown in Figure 8.21. This difference may occur in highly subjective or evaluation dimensions that require a deep understanding of design concepts, such as innovation or aesthetic value. This reflects the challenges that artificial intelligence still faces when dealing with complex and abstract evaluation tasks.

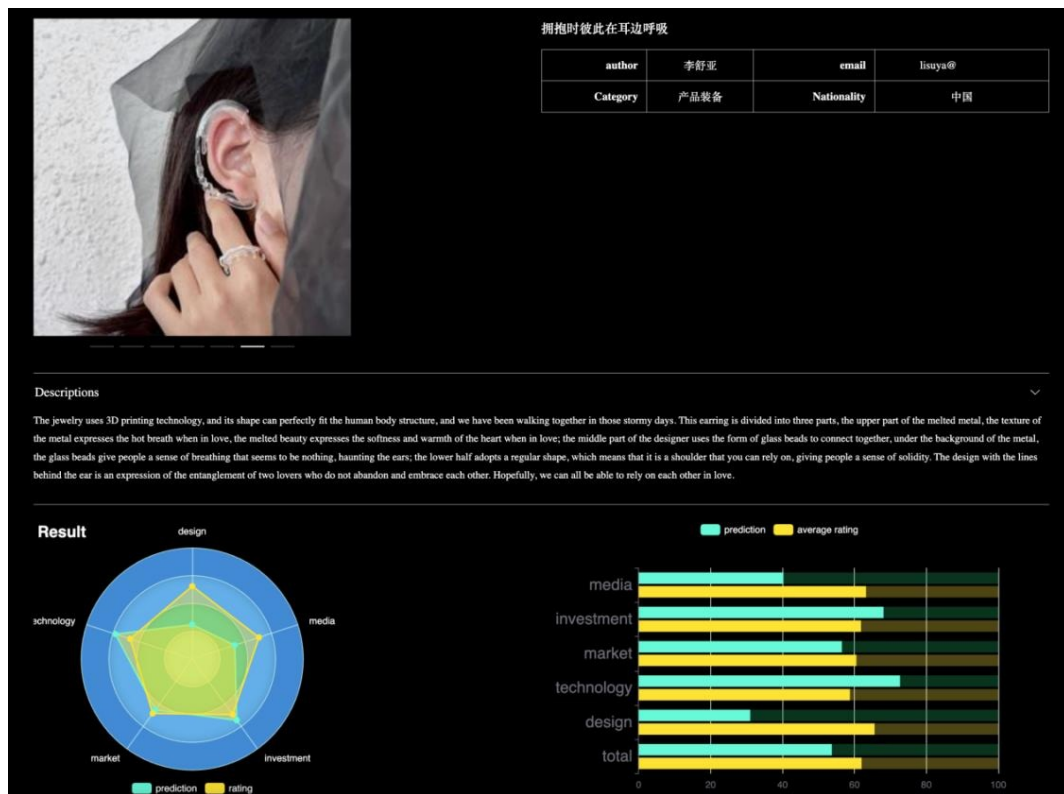


Figure 8.21 A Sample of Significant Differences

This comparative analysis not only helps to intuitively understand the current capabilities and limitations of artificial intelligence in product design evaluation but also points out the direction for future research. Researchers can conduct in-depth analyses of the reasons for deficient performance cases and improve algorithms or evaluation methods to enhance the performance of artificial intelligence in complex evaluation tasks.

8.3.2 Comparative Analysis of Human-machine Ratings

The AiDE platform adopts a multidimensional evaluation method, including five dimensions: design, technology, market, investment, and media. This multidimensional evaluation method helps to measure all aspects of product design comprehensively. The specific scoring results are shown in Table 8.1.

Table 8.1 Results of Human-machine Rating

Item	ACC	SRCC	PLCC	RMSE	MAE
design	0.3773	0.0705	0.1761	31.5657	28.8456
technology	0.7358	0.0851	0.0488	19.1562	16.4274
market	0.6603	0.2568	0.2910	22.9812	20.3050
investment	0.6226	-0.0333	-0.1006	23.4419	19.1249
media	0.6792	0.0699	0.1040	19.8676	16.6774

In terms of scoring accuracy, AI models perform differently in different dimensions. Overall, the technology dimension has the highest accuracy, reaching 73.58%, while the design dimension has the lowest accuracy, only 37.73%. This may reflect the strong subjectivity of design evaluation, and AI models still face challenges when dealing with highly subjective evaluation tasks.

Correlation analysis shows that the correlation between AI models and manual ratings is weak in most dimensions. The market dimension's Spearman rank correlation coefficient (SRCC) is the highest, at 0.2568, indicating that the ranking consistency between AI models and manual scoring is good. However, the SRCC of the investment dimension is negative (-0.0333), indicating that there may be a certain degree of contradiction between AI models and manual scoring in this dimension.

Regarding error analysis, the root means square error (RMSE) and mean absolute error (MAE) of the design dimension are the highest, at 31.5657 and 28.8456, respectively. This further confirms the difficulties AI models face in evaluating the highly subjective dimension of design. In contrast, the errors in

the technology and media dimensions are lower, indicating that AI models perform better in these objective dimensions.

It is worth noting that there are differences in the performance of human-machine comparison results when evaluating products with and without functionality. The functional product samples are shown in Figure 8.22, including electronic products, vehicles, etc. Figure 8.23 shows a sample of non-functional products. This type of product includes fashion products such as jewelry.

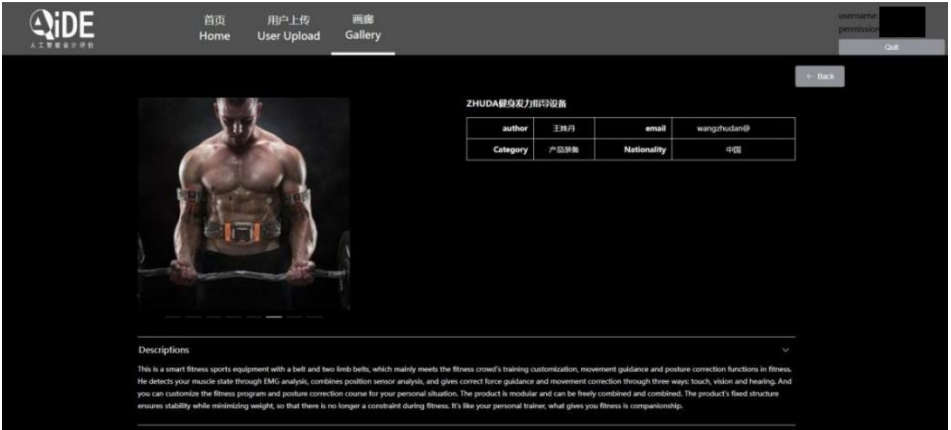


Figure 8.22 A Sample of Functional Product

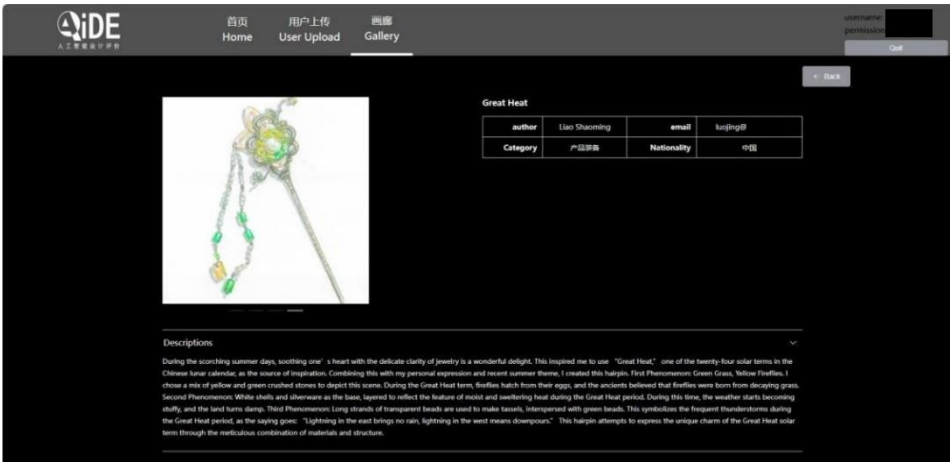


Figure 8.23 A Sample of Non-functional Product

Table 8.2 shows the comparison results of human-machine ratings for functional products. In the evaluation of functional products, the accuracy of the media dimension is as high as 85.71%.

Table 8.2 Results of Human-machine Rating for Functional Products

Item	ACC	SRCC	PLCC	RMSE	MAE
design	0.514285714	-0.21753747	-0.116970255	29.7423058	26.63186837
technology	0.714285714	0.176355232	0.177383346	21.42955329	18.91702602
market	0.714285714	-0.010505673	0.056157111	23.11983584	19.95404665
investment	0.571428571	0.00322174	-0.079814614	27.97920861	23.1542911
media	0.857142857	-0.411402157	-0.452362001	17.08956575	14.07329661

The comparison results of human-machine ratings for non-functional products are shown in Table 8.3. In evaluating products without functionality, the accuracy of this dimension is only 33.33%. This may reflect the impact of product functionality on the performance of AI model evaluation.

Table 8.3 Results of Human-machine Rating for Non-functional Products

Item	ACC	SRCC	PLCC	RMSE	MAE
design	0.111111111	0.170278638	0.240464245	35.32941189	33.54381251
technology	0.777777778	0.015479876	0.067766167	13.79906049	11.70616072
market	0.555555556	0.391124871	0.462762767	22.62252084	20.77760949
investment	0.722222222	0.131062951	0.149759655	14.86818541	12.62012955
media	0.333333333	0.327141383	0.213224408	24.38064548	21.75050756

Overall, the AI evaluation model of the AiDE platform performs well in certain dimensions, such as technology and market. However, there is still room for improvement in highly subjective dimensions, such as design. This

result reflects AI's potential and challenges in product design evaluation while emphasizing the importance of human-machine collaborative evaluation.

8.4 Artificial Intelligence Algorithm Optimization

This study conducted algorithm optimization research based on the results of the first human-machine comparison experiment. Select 60 products from the newly uploaded data, of which 44 are for training, and 16 are for testing. Of these 16 test items, eight are functional, and eight are non-functional. The purpose of this study is to explore further how to improve the performance of AI models in subjective evaluation tasks and to integrate AI and expert evaluations better to obtain more comprehensive and accurate product design evaluation results.

Figure 8.24 shows the core part of the machine learning training loop written in Python in this study. The code implements a typical deep learning training process, which includes key steps such as model training, evaluation, and saving. Specifically, the code uses a for loop to iterate through training epochs, executing the training and evaluation functions in each epoch for model training and evaluation, respectively. This design reflects the commonly used training validation pattern in machine learning, which helps monitor model performance and prevent overfitting. In addition, the pandas Data Frame was used during the algorithm optimization process to store and process the results, facilitating subsequent data analysis and visualization. This code embodies key elements in modern deep learning practices, including model training, evaluation, result recording, and model preservation, providing a foundation for improving evaluation accuracy.

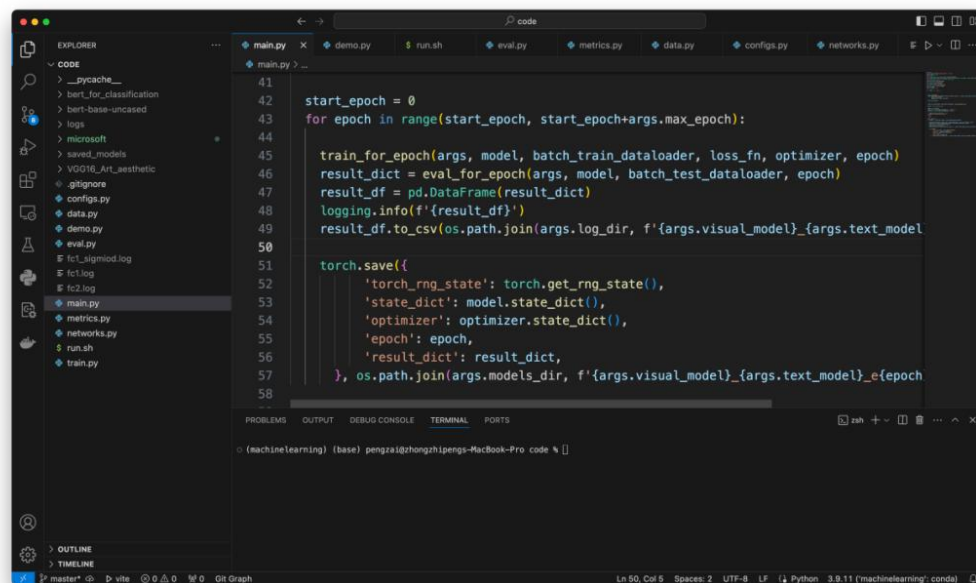


Figure 8.24 Machine Learning Training

Figure 8.25 shows the log output of the machine-learning model training process, reflecting the performance indicators of the model in multiple evaluation dimensions as they vary with training epochs. The figure contains experimental configuration information and multi-round training evaluation results, reflecting the commonly used iterative optimization process in intelligent systems. In terms of experimental setup, the log shows that the test set size is 9, and the complete training set was used for training. This configuration helps evaluate the generalization ability of the model, which is in line with the current emphasis on model reliability in intelligent manufacturing research. Performance indicators include accuracy (ACC), Spearman rank correlation coefficient (SRCC), Pearson correlation coefficient (PLC), root mean square error (RMSE), and mean absolute error (MAE). These indicators comprehensively reflect the performance of the model in different evaluation dimensions (design, technology, market, investment, and media).

```

logs > swin_bert.log
1 INFO:root:Namespace(exp_name='experiment', test_size=9, only_use_trainset=False, train
2 INFO:root:epoch 0 --- training ---
3 INFO:root:epoch 0 --- evaluating ---
4 INFO:root:      design technology      market investment      media
5 acc 1.000000 1.000000 1.000000 1.000000 1.000000
6 srcc 0.231863 -0.629817 0.303789 0.248897 0.411852
7 plcc 0.350332 -0.696847 0.118601 0.069777 0.292064
8 rmse 5.272448 9.084817 6.040406 7.603541 6.093129
9 mae 4.558433 8.597615 5.416647 6.768811 5.417953
10 INFO:root:epoch 1 --- training ---
11 INFO:root:epoch 1 --- evaluating ---
12 INFO:root:      design technology      market investment      media
13 acc 1.000000 1.000000 1.000000 1.000000 1.000000
14 srcc 0.266115 -0.535812 0.029144 0.125111 0.378173
15 plcc 0.327657 -0.713908 0.166732 -0.005985 0.207458
16 rmse 5.202055 8.825780 6.235849 6.868514 6.197581
17 mae 4.444055 8.378083 5.548768 6.134757 5.434623
18 INFO:root:epoch 2 --- training ---

```

Figure 8.25 Log Output

Figure 8.26 shows the detailed logs of the machine-learning model during the evaluation phase. This log reflects the real-time progress and performance metrics of model evaluation, which is crucial for understanding and optimizing machine learning models in intelligent manufacturing systems. The log shows that the evaluation process is divided into 14 steps, and each step's completion percentage and time consumption are accurately recorded. This fine-grained progress tracking helps researchers monitor the efficiency of model evaluation and identify potential performance bottlenecks. The processing speed of each evaluation step (in samples per second) varies as the evaluation process progresses, gradually increasing from an initial 16.17 samples per second to a final 4.71 samples per second. This speed variation reflects certain data batch processing or model computation characteristics, providing clues for further optimizing the evaluation process. At the end of the log, specific performance indicators of the model in the "design" dimension are also included, such as accuracy (ACC), Spearman rank

correlation coefficient (SRCC), and Pearson correlation coefficient (PLCC). The accurate recording of these indicators is crucial for evaluating the model's performance in product design evaluation tasks. The performance of models varies in different dimensions. For example, in the technology dimension, SRCC and PLCC show negative values, which may reflect the dimension evaluation's complexity or the data features' specificity.

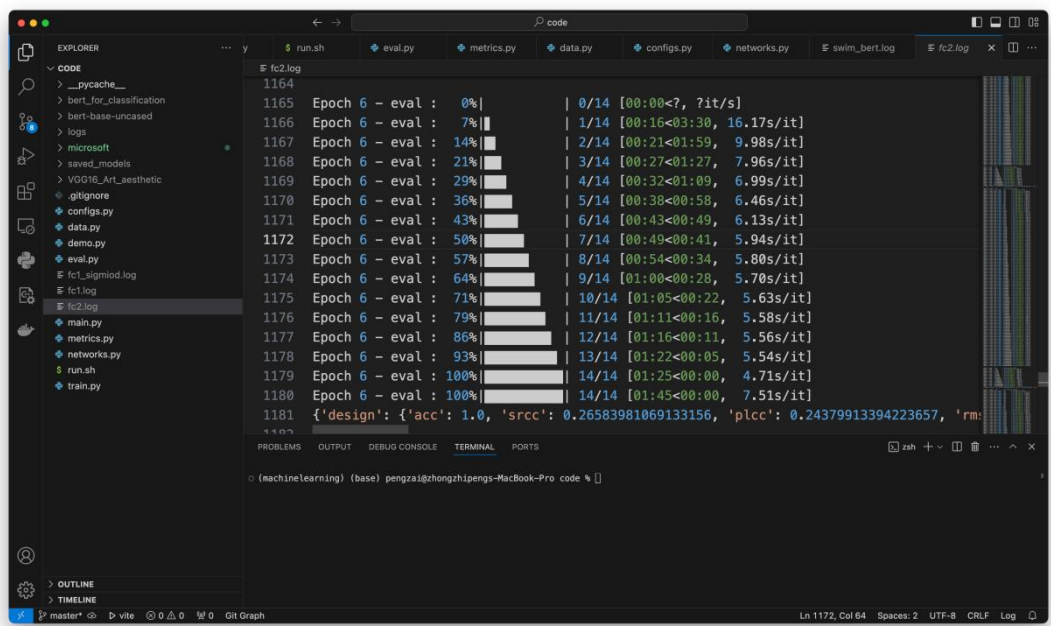


Figure 8.26 Training and Validation

This study provides an important basis for optimizing model performance and understanding the relationship between evaluation indicators through multiple training rounds and multidimensional evaluation. The optimized test metrics are shown in Table 8.4. The optimized AI model shows significant performance improvements in multiple evaluation dimensions. The improvement is particularly significant in the three indicators of accuracy (ACC), root mean square error (RMSE), and mean absolute error

(MAE). This comprehensive performance improvement reflects the effectiveness of the model optimization strategy.

Table 8.4 Optimized Human-machine Rating Results

Item	ACC	SRCC	PLCC	RMSE	MAE
design	0.823529412	0.452483224	0.581900317	8.895765009	7.697040846
technology	0.705882353	0.598039216	0.581334071	9.37186168	8.03707503
market	0.647058824	0.431372549	0.447508422	9.817422569	7.90984256
investment	0.764705882	0.541666667	0.487416381	9.843799728	8.279466702
media	0.705882353	0.652360638	0.647872418	7.24637821	6.538562864

The accuracy of the design dimension has significantly increased from 37.73% before optimization to 82.35%, RMSE has decreased from 31.5657 to 8.8958, and MAE has decreased from 28.8456 to 7.6970. This significant improvement indicates that the optimized model manages highly subjective design evaluation tasks better. Regarding technology, although the accuracy has slightly decreased (from 73.58% to 70.59%), both SRCC and PLCC have improved. This indicates that the optimized model performs better in rating ranking and linear correlation, although slightly reduced in binary classification tasks.

In addition, Table 8.5 and Table 8.6 indicate that the optimized model performs well in functional and non-functional product evaluations. Especially in the evaluation of non-functional products, the accuracy of all dimensions reaches or exceeds 83.33%, and SRCC and PLCC are higher than 0.8. This result indicates that the model can be generalized when dealing with different product evaluation tasks. However, the research results also showed significant negative SRCC and PLCC in certain dimensions, such as the

technology dimension, indicating that there may be an inverse relationship between the model's rating prediction in these dimensions.

Table 8.5 Optimized Results for Functional Products

Item	ACC	SRCC	PLCC	RMSE	MAE
design	0.904761905	0.678659698	0.65974527	6.133617631	4.611158803
technology	0.833333333	0.645247549	0.640607763	6.635089331	5.156353716
market	0.857142857	0.695919939	0.655885887	6.609763854	4.521013239
investment	0.80952381	0.643815081	0.604114141	6.896801554	5.163729377
media	0.833333333	0.700700945	0.709282884	4.956727126	3.882701165

Table 8.6 Optimized Results for Non-functional Products

Item	ACC	SRCC	PLCC	RMSE	MAE
design	0.833333333	0.91744066	0.867148493	3.872711372	3.257800789
technology	0.888888889	0.933952528	0.918987614	4.159852941	3.532600358
market	0.833333333	0.820433437	0.848888446	4.241288242	3.308385216
investment	0.888888889	0.882352941	0.87585221	3.842327161	2.967936576
media	0.944444444	0.872033024	0.869814857	3.398832723	2.763159144

The optimized AI model has shown significant performance improvements on the AiDE platform, particularly in accuracy, error control, and correlation. These improvements enhance the model's reliability in product design evaluation tasks and provide dedicated support for further optimization and application. Future research can delve into the performance differences of models in specific dimensions to further enhance the overall effectiveness of AI-assisted product design evaluation.

8.5 Summary and Discussions

In design evaluation experiments, research increasingly focuses on the collaboration between humans and artificial intelligence (Chong et al., 2022). The study emphasizes the importance of human designers effectively utilizing artificial intelligence input to achieve success (Chong et al., 2022). Human abilities and confidence are crucial in working with artificial intelligence systems (Chong et al., 2022). Challenges may arise when integrating non-human agents such as AI into the design process and teams, highlighting the need to address key issues of seamless human-AI collaboration (Figoli et al., 2022).

Jiang et al. (2022) proposed the need for a comprehensive framework to combine artificial intelligence capabilities with context-aware techniques to enhance collaborative performance. This study is crucial for understanding the role of context in design evaluation experiments involving artificial intelligence and human designers (Jiang et al., 2022). Rezwana's (2022) research provides insights into user perception and collaborative experience in human-computer interaction design, providing valuable knowledge for developing effective co-creation systems in various fields. The ethical considerations of co-creation between humans and artificial intelligence have been discussed. Rezwana and Maher (2022) identified and discussed the ethical issues that may arise when artificial intelligence becomes a partner in collaborative efforts. They believe ethical implications must be carefully considered to ensure responsible and effective design practices in human-AI collaboration (Rezwana & Maher, 2022). Verganti et al. (2020) have explored the transformative impact of artificial intelligence on innovation and design processes, emphasizing the user-centered nature, creativity, and continuous learning capabilities that AI brings to the design field.

A study introduced an evaluation model based on fuzzy dual experience. This model combines engineering design with customer response, demonstrating how this approach enhances the product development experience and creates value through customer engagement (Chen, 2016). It provides a structured framework for evaluating design solutions that meet engineering requirements and customer preferences, highlighting the importance of a comprehensive approach in design evaluation (Chen, 2016). The participatory meaning construction in abstract painting has been studied, which can provide research references for the design of collaborative systems involving artificial intelligence and human creators (Davis et al., 2016). A collaborative framework for human-AI interaction in data-driven text generation has been proposed, emphasizing the importance of designing machine learning systems that prioritize end-user control scenarios in model development (Strobelt et al., 2022).

The human-centered design approach in artificial intelligence systems is crucial for ensuring effective collaboration and user satisfaction (Strobelt et al., 2022). Karimi et al. (2018) proposed a framework for evaluating creativity in systems where computers and human users collaborate to complete creative tasks, exploring the issue of creativity assessment in computational co-creation systems. Linsey et al. (2007) have studied the impact of analogical product descriptions on designers' problem-solving abilities, demonstrating the influence of language and representation on design outcomes. A study delved into the neural responses of experiential product design and its relationship with personalized preferences, providing insights for product design evaluation based on personal preferences and neural processes (Ma et al., 2018). The proposal of a collaborative learning model

between humans and artificial intelligence in urban search and rescue missions demonstrates the potential of AI to improve learning and performance outcomes in complex tasks (Schoonderwoerd et al., 2022).

The design evaluation experiment's comprehensive collaboration between humans and artificial intelligence highlights the intricate interplay between human creativity, artificial intelligence capabilities, and ethical considerations. By utilizing insights from research exploring the dynamic interaction between humans and artificial intelligence, researchers in the design field can develop algorithmic models to evaluate design solutions. These design solutions leverage the advantages of human designers and artificial intelligence systems. This collaborative approach improves the efficiency and effectiveness of the design process, providing a feasible pathway for integrating AI-driven innovation with ethical design practices. This method not only improves the efficiency of the design process but also provides new ways of addressing ethical and social issues brought about by artificial intelligence.

This study explores the potential application of artificial intelligence in product design evaluation by developing and implementing the AiDE platform. The experimental results reveal the advantages and limitations of AI models in design evaluation tasks, highlighting the importance of human-machine collaboration in improving evaluation quality. The user interface design of the AiDE platform adopts a concise and clear style, effectively reducing visual interference and improving user experience. The design of multimodal data input and visualized scoring results reflects the latest trends in product design evaluation. However, the potential impact of interface design on manual scoring still needs further research. For example,

although the sliding bar scoring method allows precise scoring, it may lead raters to focus too much on numerical values and ignore the overall impression. Future research can explore how to improve the accuracy and consistency of scoring by improving GUI design, such as introducing more intuitive scoring methods or providing richer reference information.

The experimental results demonstrate the potential of human-machine collaboration in product design evaluation research. AI models perform well in specific dimensions, such as technology and market, but there is still room for improvement in highly subjective dimensions, such as design. This difference reflects the advantages of AI models in handling objective indicators and the limitations of understanding abstract concepts and innovation. The human-machine collaboration evaluation method has the potential to bridge this gap by combining the high efficiency of AI with the insights of human experts. Future research should delve into the learning mechanism of AI models, clarify the boundaries of their evaluation capabilities, and study how to effectively integrate AI scoring and manual scoring to optimize the product design evaluation process.

This study provides important theoretical contributions to artificial intelligence-assisted product design evaluation. Firstly, it demonstrates the potential of multimodal deep learning in design evaluation, providing innovative ideas for building more comprehensive and accurate evaluation models. Secondly, by comparing the rating results of AI models and human experts, this study reveals the complexity and necessity of human-machine collaboration in design evaluation, laying the foundation for establishing a more effective human-machine collaboration evaluation framework. Finally, the evaluation method and indicator system proposed in this study provide

reliable reference standards for future research. Future research directions include the following points.

- Expanding sample size and diversity to improve the model's generalization ability.
- Optimize website functionality by introducing more interpretable human-computer interaction designs to enhance user experience.
- Deeply exploring the interpretability of AI models and improving evaluation results' credibility.
- Research how to integrate AI and human scoring more effectively, leveraging their strengths.
- Exploring the long-term impact of AI-assisted evaluation on design education and practice.

This study also has limitations. Firstly, the limited sample size may affect the generalizability of the results. Secondly, there is room for optimization in GUI design and website functionality, such as improving system response speed and adding more interactive features. In addition, the performance of AI models in handling highly subjective evaluation dimensions still needs improvement. However, the results of Study 4 provide an important foundation for the next stage of research (Study 5). Study 5 will focus on

verifying the practical application value of AI-assisted evaluation tools and exploring designers' acceptance of AI evaluation results.

In summary, this study provides new insights and methods for applying artificial intelligence in product design evaluation through developing and implementing the AiDE platform. It reflects the latest trends in product design evaluation. It provides rich directions and ideas for future research, which is expected to promote the progress of design evaluation methods and improve design quality.

Study 4 verifies the evaluation performance of the AiDE and explores the possibility of human-machine collaborative evaluation by designing and implementing a human-machine comparative evaluation experiment. The experimental results verify the AI model's effectiveness and provide important input for the prototype verification study (Study 5). Although the experiment may have certain limitations, its results offer key insights into understanding the application value of AI in actual design evaluation. Study 4 lays the foundation for the prototype verification study in the next chapter, making it possible to explore AI tools' applications in actual design environments further.

Chapter 9. Study 5- Prototype Validation Research

This chapter introduces a prototype validation study that aims to delve into the user experience and satisfaction of the AiDE. Through a mixed method (questionnaire and structured interviews), Study 5 takes a deep dive into user experience and satisfaction with the AI-assisted design evaluation tool. By collecting user feedback on the model's prediction results, Study5 aims to gain a more comprehensive understanding of the system's strengths and limitations, providing an important basis for future optimization and improvement while evaluating the system's application value and acceptance in the actual product design process. The research results of this chapter will lay a solid foundation for the further improvement and promotion of the AiDE system and provide important empirical references for the application of artificial intelligence in product design evaluation.

9.1 Introduction

This study aims to develop and validate an AI-assisted product design evaluation system through in-depth exploratory research. In the preliminary research, researchers identified the needs of designers for AI tools (Study 1), constructed a multimodal product design dataset (Study 2), and designed and implemented an AI evaluation model based on deep learning (Study 3). Study 4 evaluated the accuracy and reliability of the AI model through human-computer comparison experiments, and the results showed that the AI model performed well in certain evaluation dimensions. However, there is still room for improvement in highly subjective dimensions. In artificial intelligence (AI), it is crucial to ensure the effectiveness and usability of AI

models before deploying them in real-world scenarios. A key aspect of this process is prototype validation. It involves rigorous evaluation to ensure the reliability and generality of artificial intelligence models (Levman et al., 2023).

Anggraini (2022) emphasized that expert validation is crucial in the prototype validation process. Expert review provides valuable insights into the design and functionality of artificial intelligence models, helping to identify potential weaknesses, inaccuracies, or biases that may be overlooked during the development phase. Prototype evaluation by domain experts affords developers holistic insights into design merits and constraints, thereby facilitating data-driven refinement strategies. Li (2023) focuses on the usability evaluation and enhancement of artificial intelligence frameworks, emphasizing user-centered design principles' importance in optimizing AI features and interactions. By prioritizing user feedback and incorporating usability evaluation strategies into the development process, developers can create AI systems that are intuitive, user-friendly, and meet user expectations and preferences. This user-centered approach to artificial intelligence design and evaluation helps ensure that AI technology meets the needs and requirements of its target users, improving its usability and acceptance.

As emphasized by Cerqueira et al. in their 2022 work, ethical considerations also play a significant role in the development and evaluation of artificial intelligence (Cerqueira et al., 2022). By adopting design science research methods and incorporating ethical needs into the prototype development process, researchers can ensure that artificial intelligence systems adhere to ethical guidelines and principles. Actively addressing ethical issues in artificial intelligence development can help mitigate potential risks while

promoting responsible innovation, thereby avoiding unexpected outcomes and ethical dilemmas. Alufaisan et al. (2020) delved into the impact of explainable artificial intelligence on human decision-making, emphasizing the importance of interpretability, trustworthiness, and availability in artificial intelligence systems. By enhancing the interpretability of AI models and ensuring that users can understand and trust the decisions made by these systems, developers can improve user acceptance and decision outcomes. This focus on interpretability and usability highlights the crucial role of transparency and user-friendliness in artificial intelligence systems, particularly in human decision-making applications.

In summary, prototype validation and usability evaluation are critical to developing artificial intelligence models and systems. Developers can create dependable, effective, and user-friendly AI technologies by following strict validation processes, integrating expert feedback, and prioritizing user-centered design principles. Ethical considerations, usability standards, and interdisciplinary collaboration play a crucial role in ensuring that artificial intelligence systems meet the needs and expectations of users in various fields. Through comprehensive evaluation methods and a focus on usability, artificial intelligence developers can improve the availability and effectiveness of AI systems, contributing to their successful deployment and adoption in practical applications.

Based on the findings of previous research (Study 1-Study 4), Study 5 aims to validate further the prototype of the AI-assisted product design evaluation system and gain a deeper understanding of user experience and satisfaction. This research stage is crucial for evaluating the system's practical application value and acceptance. By collecting user feedback on the model's prediction

results, this study can gain a more comprehensive understanding of the system's strengths and limitations, providing an important basis for future optimization and improvement.

9.2 Methods

Study 5 adopts a mixed research approach in this validation phase, combining questionnaire surveys and structured interviews. This method can collect quantitative data to evaluate the overall performance of the system, as well as explore users' subjective experiences and insights through qualitative data. Questionnaire surveys can obtain users' overall evaluations of the system. At the same time, structured interviews can capture more subtle user feedback and suggestions, which helps to understand users' needs and expectations. Through this comprehensive verification method, this study aims to comprehensively evaluate the practicality and effectiveness of AI-assisted product design evaluation systems, laying a solid foundation for future system optimization and promotion applications.

9.2.1 Questionnaire

This study distributed questionnaires to 129 users who participated in the human-machine comparison experiment in the previous chapter and had experience using the AiDE platform. The questionnaire consists of Likert scale scoring questions and open-ended questions. This structural design allows researchers to collect quantitative and qualitative data to fully understand designers' attitudes and opinions on AI-assisted design evaluation tools. The questionnaire uses a 5-point Likert scale ranging from -2 to 2. This design allows respondents to express different degrees of agreement or disagreement (Fujii et al., 2020).

The questionnaire contains the following main parts: AI Evaluation Trust Scale, AI Evaluation Tool Use Willingness Scale, and open-ended question design. The AI Evaluation Trust Scale contains ten questions designed to assess designers' trust in the application of AI in conceptual design evaluation. The questions cover aspects such as familiarity with AI, willingness to use, and views on AI accuracy. The AI Evaluation Tool Use Willingness Scale contains seven questions that assess designers' willingness to use AI evaluation tools. The questions cover time efficiency, innovation, and willingness to learn new tools. The open-ended questions contain five open-ended questions designed to collect designers' in-depth insights and suggestions on the application of AI in design evaluation. The design of open-ended questions allows respondents to express their ideas freely, providing rich qualitative data.

9.2.2 Structured Interview

This study adopts a structured interview method to explore the application and impact of AI-assisted design evaluation tools in depth. Ten interviewees were randomly selected from 129 users of the AiDE system, including five junior product designers and five design experts. Each interview lasted about 1 hour. The interview outline covers nine main parts, which can be categorized into three core research directions.

The first research direction focuses on the role and application of AI in design evaluation. This includes exploring AI's role in conceptual design evaluation, how AI can be integrated into design education, and the ideal human-computer collaboration model. Through these discussions, researchers

can better understand how AI can be integrated into the design process and maximize its potential, thereby providing new insights and methods for product and engineering design.

The second research direction focuses on improving and optimizing AI evaluation tools. This includes exploring how to improve the trust and willingness to use AI evaluation tools, improving AI evaluation functions, and the focus of future research. These discussions are crucial to improving the practicality and acceptance of AI tools and provide valuable guidance for the future application of AI in design.

The third research direction explores the ethical and social impact of AI applications. This includes considering the ethical and privacy issues that may arise from the application of AI in design evaluation, as well as its broader impact on the design industry and society. Through the introduction and conclusion, this study can put these discussions in a broader context and summarize the key findings, providing clear research directions for the future application of AI in industrial design.

This structured interview method and classification framework allow researchers to systematically collect and analyze qualitative data and gain an in-depth understanding of users' experiences and opinions, providing valuable insights for further improving AI-assisted design evaluation tools. By comprehensively considering the role of AI in design, tool optimization, and social impact, this study provides comprehensive theoretical and practical guidance for the future application of AI in product design and engineering design.

9.3 Data Analysis and Results

9.3.1 Questionnaire Analysis Results

This study used various statistical methods to analyze the questionnaire data, including descriptive statistics, reliability analysis, correlation analysis, etc., to evaluate designers' attitudes and acceptance of AI comprehensively assisted design evaluation tools. The study used Cronbach's Alpha coefficient to verify the scale's internal consistency (Lu et al., 2023). In terms of descriptive statistics, the study analyzed each item's mean, standard deviation, and median. The study also conducted extensive correlation analysis. In general, this study comprehensively evaluated designers' attitudes and acceptance of AI-assisted design evaluation tools through a combination of quantitative and qualitative methods, providing important empirical evidence for future AI tools development and promotion. The specific research results will be presented in the following sections.

9.3.1.1 The Limitations of AI Design Evaluation

Designers encounter key issues when using the AiDE platform, as shown in Figure 9.1. Insufficient or low-quality data, lack of trust in AI evaluation results, lack of industry standards and specifications, and insufficient understanding of AI algorithm operations are considered major obstacles.

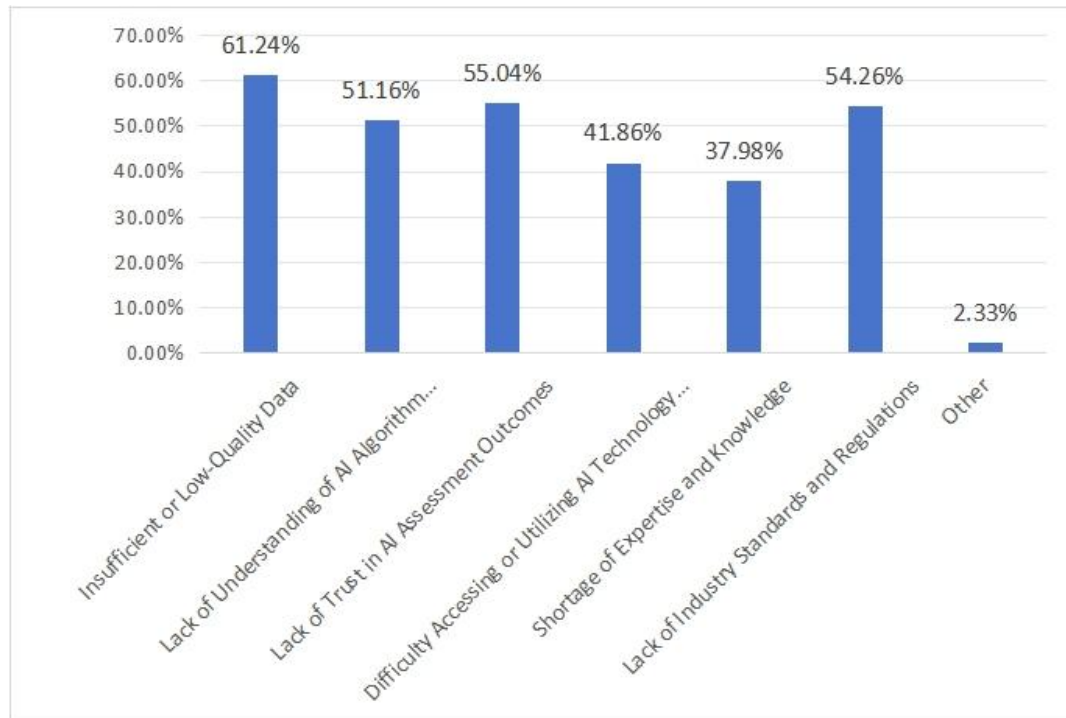


Figure 9.1 Main Limitations of AI in Concept Evaluation

According to the survey results in Table 9.1, 61.24% of designers believe that insufficient data quality is the main problem, followed by a lack of trust in AI evaluation results (55.04%) and a lack of industry standards and specifications (54.26%). In addition, more than half of designers (51.16%) also expressed concerns about insufficient understanding of AI algorithm operation. Relatively speaking, although more common, difficulties in accessing or utilizing AI technology and tools (41.86%) and shortage of professional knowledge and skills (37.98%) appear less prominent in comparison. The prevalence of these challenges indicates that in promoting the widespread application of AI technology in design evaluation in the future, data quality, enhancing trust, improving industry standards, and improving understanding of AI operation must be addressed first, as these factors are key to improving designer acceptance and efficiency.

Table 9.1 Summary of Response Rate and Penetration Rate

project	Response		Popularity ($n=129$)
	n	Response Rate	
Insufficient or Low-Quality Data	79	20.15%	61.24%
Lack of Understanding of AI Algorithm Operation	66	16.84%	51.16%
Lack of Trust in AI Assessment Outcomes	71	18.11%	55.04%
Difficulty Accessing or Utilizing AI Technology and Tools	54	13.78%	41.86%
Shortage of Expertise and Knowledge	49	12.50%	37.98%
Lack of Industry Standards and Regulations	70	17.86%	54.26%
Other	3	0.77%	2.33%
Project	392	100%	303.88%

Note: Goodness of fit test $\chi^2 = 69.857$ $p = 0.000$

Table 9.2 shows the co-occurrence of various barriers designers encounter when using AI for conceptual design evaluation, revealing the inter-correlation between these factors. The data show that insufficient or low-quality data, insufficient understanding of the operation of AI algorithms, lack of trust in AI evaluation results, difficulty in accessing technology, shortage of expertise and skills, and lack of industry standards and specifications are not isolated issues but are intertwined. For example, the problem of insufficient or low-quality data is not only related to an insufficient understanding of the operation of AI algorithms but also closely related to a lack of trust in AI evaluation results. It also has a high co-occurrence rate, difficulties accessing technology, and a need for industry

standards. This multifaceted challenge shows that to promote AI application in design evaluation effectively, it is necessary to comprehensively consider and improve multiple aspects, such as education and training, technology development, and industry specification formulation, to improve designers' trust in AI tools and their use efficiency.

Table 9.2 Option Co-occurrence Matrix

project	Insufficient or Low-Quality Data	Lack of Understanding of AI Algorithm Operation	Lack of Trust in AI Assessment Outcomes	Difficulty Accessing or Utilizing AI Technology and Tools	Shortage of Expertise and Knowledge	Lack of Industry Standards and Regulations	Other
Insufficient or Low-Quality Data	-	-	-	-	-	-	-
Lack of Understanding of AI Algorithm Operation	37	-	-	-	-	-	-
Lack of Trust in AI Assessment Outcomes	45	43	-	-	-	-	-
Difficulty Accessing or Utilizing AI Technology and Tools	34	33	32	-	-	-	-
Shortage	28	26	24	22	-	-	-

of Expertise and Knowledge							
Lack of Industry Standards and Regulations	44	34	34	30	34	-	-
Other	2	0	0	0	1	2	-

This result shows that improvements must be made from multiple angles when promoting AI-assisted design evaluation tools, including increasing transparency, improving data quality, simplifying technical tools, strengthening education and training, and establishing industry standards.

9.3.1.2 Trust in AI Evaluation Models and Willingness to Use

Verified by Cronbach's Alpha coefficient, the reliability coefficient of the AI evaluation trust scale is 0.843, as shown in Table 9.3. This indicates that the data has high internal consistency.

Table 9.3 AI Evaluation Trust Scale Cronbach Reliability Analysis

Number of items	Sample size	Cronbach α coefficient
10	129	0.843

($\alpha > 0.8$, high reliability; $0.7 < \alpha < 0.8$, good reliability; $0.6 < \alpha < 0.7$, acceptable reliability; $\alpha < 0.6$, poor reliability)

The analysis results of Table 9.4 show that designers have a cheerful outlook toward applying AI in conceptual design evaluation, but there are significant

differences in their trust in its accuracy. It is particularly noteworthy that designers are less familiar with AI (Question 1, mean value -0.178), and their confidence in AI providing accurate results needs to be improved (Questions 4 and 7, mean values are 0.372 and 0.411, respectively).

Table 9.4 AI Evaluation Trustworthiness Rating

topic	Sample size	Minimum value	Maximum value	Mean	SD	Median
1.How familiar are you with using AI in conceptual design evaluation?	129	-2.000	2.000	-0.178	0.947	0.000
2.Are you willing to use AI algorithms for quantitative evaluation of your designs?	129	-2.000	2.000	0.907	0.744	1.000
3.What do you think about using AI for scoring or evaluating conceptual designs?	129	-2.000	2.000	0.705	0.795	1.000
4.Do you believe AI can provide accurate results in conceptual design evaluation?	129	-2.000	2.000	0.372	0.761	0.000
5.Can AI technology effectively assist your design decision-making process?	129	-1.000	2.000	0.736	0.702	1.000
6.Are you satisfied with AI's suggestions for conceptual design evaluation?	129	-2.000	2.000	0.442	0.672	0.000
7.Do you trust AI's accuracy and usefulness in conceptual design evaluation?	129	-2.000	2.000	0.411	0.756	0.000
8.Do you trust AI's accuracy and usefulness	129	-2.000	2.000	0.597	0.644	1.000

in the optimization of conceptual designs?						
9.Are you open to using AI technology in conceptual design evaluation?	129	-2.000	2.000	0.915	0.696	1.000
10.I am willing to rely on AI's evaluation to revise my designs.	129	-2.000	2.000	0.527	0.811	1.000

The Cronbach's Alpha coefficient of the AI evaluation tool willingness scale is 0.752, as shown in Table 9.5, which also shows good reliability.

Table 9.5 Cronbach Reliability Analysis Using the Willingness Scale

Number of items	sample size	Cronbach α coefficient
7	129	0.752

($\alpha > 0.8$, high reliability; $0.7 < \alpha < 0.8$, good reliability; $0.6 < \alpha < 0.7$, acceptable reliability; $\alpha < 0.6$, poor reliability)

Table 9.6 shows that designers are inclined to use AI evaluation tools, especially the potential of AI in improving efficiency and innovation (questions 11 and 14, mean values 1.202 and 1.062, respectively). However, designers hope to have a deeper understanding of AI tools before using them (question 12, mean value 0.589). They may be reserved when AI suggestions do not match their ideas (question 47, mean value 0.411).

Table 9.6 Willingness to Use AI Evaluation Tools

Topic	Sample size	Minimum value	Maximum value	Mean	SD	Median
11.Would you use AI for conceptual design evaluation if it saves time?	129	-1.000	2.000	1.202	0.733	1.000

12.Would you use AI in conceptual design evaluation without fully understanding its workings?	129	-2.000	2.000	0.589	0.880	1.000
13.Would you prefer AI if its suggestions increase design innovation?	129	-1.000	2.000	0.984	0.625	1.000
14.Are you willing to learn new tools to harness AI's benefits in design evaluation?	129	-1.000	2.000	1.062	0.622	1.000
15.Do you believe AI improves the market performance of final products in design evaluation?	129	-2.000	2.000	0.674	0.762	1.000
16.Are you willing to share design data to improve AI evaluation accuracy?	129	-2.000	2.000	0.620	0.894	1.000
17.Would you consider AI suggestions even if they sometimes differ from your ideas?	129	-2.000	2.000	0.411	0.777	1.000

The correlation analysis in Table 9.7 shows that except for the degree of understanding of AI, the other trust evaluation questions are significantly positively correlated with the willingness to use ($p < 0.01$). This shows that trust is a key factor in promoting the willingness to use and improving designers' trust in AI tools can effectively promote the application and popularization of these tools.

Table 9.7 Pearson Correlation of Each Dimension

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Q	-0.12	0.407*	0.291*	0.201*	0.302*	0.341*	0.230*	0.273*	0.524*	0.280*

11	8	*	*		*	*	*	*	*	*
Q 12	0.005	0.466* *	0.384* *	0.381* *	0.380* *	0.349* *	0.314* *	0.353* *	0.401* *	0.459* *
Q 13	-0.15 0	0.383* *	0.211* *	0.160	0.436* *	0.240* *	0.195*	0.315* *	0.482* *	0.340* *
Q 14	-0.12 7	0.333* *	0.211* *	0.231* *	0.396* *	0.121	0.195*	0.258* *	0.481* *	0.337* *
Q 15	0.006	0.525* *	0.202* *	0.440* *	0.481* *	0.283* *	0.410* *	0.463* *	0.433* *	0.432* *
Q 16	0.085	0.369* *	0.171	0.381* *	0.374* *	0.346* *	0.406* *	0.343* *	0.387* *	0.246* *
Q 17	-0.03 8	0.472* *	0.223* *	0.427* *	0.329* *	0.353* *	0.296* *	0.349* *	0.296* *	0.373* *

* $p < 0.05$ ** $p < 0.01$, Q (Question)

The results show that although designers have a cheerful outlook towards AI applications, there are significant individual differences. Both the trust scale and the willingness to use scale showed good internal consistency reliability (Cronbach's Alpha was 0.843 and 0.752, respectively). Correlation analysis revealed a significant positive correlation between trust and willingness to use. However, there needed to be a meaningful relationship between designers' familiarity with AI and their willingness to use it. The study also emphasized the importance of transparency in AI systems and the need to balance AI assistance and human creativity in human-machine collaboration.

9.3.1.3 Relationship between Familiarity with AI and Confidence/

Willingness in Using AI

As shown in Table 9.8 and Table 9.9, the study found no significant relationship between designers' familiarity with AI and their confidence and willingness to use AI. Except for Question 1 ($r=0.266$, $p < 0.01$) and Question 14

($r=0.165$), the correlation coefficients of most questions were low. This suggests that more than simply increasing familiarity with AI may be required to enhance designers' confidence and willingness to use AI.

Table 9.8 The Relationship between Familiarity and Using Confidence

	Familiarity with the application of AI in design
1.How familiar are you with using AI in conceptual design evaluation?	0.266**
2.Are you willing to use AI algorithms for quantitative evaluation of your designs?	0.107
3.What do you think about using AI for scoring or evaluating conceptual designs?	0.115
4.Do you believe AI can provide accurate results in conceptual design evaluation?	0.045
5.Can AI technology effectively assist your design decision-making process?	0.075
6.Are you satisfied with AI's suggestions for conceptual design evaluation?	0.143
7.Do you trust AI's accuracy and usefulness in conceptual design evaluation?	0.015
8.Do you trust AI's accuracy and usefulness in the optimization of conceptual designs?	0.111
9.Are you open to using AI technology in conceptual design evaluation?	0.042
10.I am willing to rely on AI's evaluation to revise my designs.	0.118
* $p<0.05$ ** $p<0.01$	

Table 9.9 The Relationship between Familiarity and Using Willingness

	Familiarity with the application of AI in design
11.Would you use AI for conceptual design evaluation if it saves time?	-0.003

12.Would you use AI in conceptual design evaluation without fully understanding its workings?	0.095
13.Would you prefer AI if its suggestions increase design innovation?	0.070
14.Are you willing to learn new tools to harness AI's benefits in design evaluation?	0.165
15.Do you believe AI improves the market performance of final products in design evaluation?	0.072
16.Are you willing to share design data to improve AI evaluation accuracy?	0.029
17.Would you consider AI suggestions even if they sometimes differ from your ideas?	0.052
* $p<0.05$ ** $p<0.01$	

This study explored the Relationship between designers' familiarity with artificial intelligence (AI) and their confidence and willingness to use it. The study found no significant correlation between designers' familiarity with AI and their confidence and willingness to use it. Except for the familiarity with using AI for conceptual design evaluation, which showed a weak positive correlation ($r=0.266$, $p<0.01$), the correlation coefficients of other questions were primarily low. In terms of willingness to use AI, except for willingness to learn new tools to take advantage of AI, which showed a slight positive correlation ($r=0.165$), the correlation coefficients of the remaining questions were close to zero. This finding challenges the common assumption that familiarity directly affects willingness to use, suggesting that simply increasing familiarity with AI may not enhance designers' confidence and willingness to use AI.

9.3.1.4 Relationship between Length of Time Engaged in Design and Confidence in Using AI

The analysis results of Tables 9.10 and Table 9.11 show that the relationship between designers' work experience in the design industry and their confidence and willingness to use AI is not strong. The correlation coefficients of most questions are close to zero or exceedingly small, which means that design experience may not be the main factor affecting the acceptance of AI tools.

Table 9.10 The Relationship between Working Experience and Using Confidence

	Duration of working in the design industry
1.How familiar are you with using AI in conceptual design evaluation?	0.024
2.Are you willing to use AI algorithms for quantitative evaluation of your designs?	-0.040
3.What do you think about using AI for scoring or evaluating conceptual designs?	-0.138
4.Do you believe AI can provide accurate results in conceptual design evaluation?	0.080
5.Can AI technology effectively assist your design decision-making process?	-0.088
6.Are you satisfied with AI's suggestions for conceptual design evaluation?	-0.051
7.Do you trust AI's accuracy and usefulness in conceptual design evaluation?	0.015
8.Do you trust AI's accuracy and usefulness in the optimization of conceptual designs?	0.099
9.Are you open to using AI technology in conceptual design evaluation?	-0.038
10.I am willing to rely on AI's evaluation to revise my designs.	-0.027

* $p<0.05$ ** $p<0.01$

Table 9.11 The Relationship between Working Experience and Using
Willingness

	Duration of working in the design industry
11.Would you use AI for conceptual design evaluation if it saves time?	0.070
12.Would you use AI in conceptual design evaluation without fully understanding its workings?	0.085
13.Would you prefer AI if its suggestions increase design innovation?	0.048
14.Are you willing to learn new tools to harness AI's benefits in design evaluation?	0.109
15.Do you believe AI improves the market performance of final products in design evaluation?	-0.018
16.Are you willing to share design data to improve AI evaluation accuracy?	0.001
17.Would you consider AI suggestions even if they sometimes differ from your ideas?	0.080
* $p<0.05$ ** $p<0.01$	

The study found no significant correlation between designers' work experience in the design industry and their confidence and willingness to use AI. The correlation coefficients for most questions were close to zero or exceedingly small, indicating that the Relationship between work experience and AI acceptance was weak. Regarding willingness to use AI, only willingness to learn new tools to take advantage of AI showed a weak positive correlation ($r=0.109$), and the correlation coefficients of the remaining

questions were low. This finding challenges the common assumption that work experience affects technology acceptance, suggesting that other factors, such as the availability, transparency, and integration of AI systems with existing workflows, may influence designers' acceptance of AI. This result emphasizes the need to consider a broader range of factors when promoting AI-assisted design tools rather than relying on designers' work experience. Future research should further explore the key factors that influence designers' acceptance of AI technology and how to integrate AI more effectively into the work practices of designers with dissimilar experience levels.

9.3.1.5 Open Questions

Through open-ended questions, this study investigated designers' views, expectations, and suggestions on artificial intelligence (AI) in conceptual design evaluation. The survey covered the role of AI in design evaluation, key functions, areas for improvement, application improvement suggestions, and future development prospects. These questions aim to gain a deeper understanding of designers' cognition, needs, and attitudes toward AI-assisted design evaluation tools, providing an important reference for developing and improving future AI systems.

One of the research results shows that designers believe that AI should support conceptual design evaluation involving data collection and analysis, idea generation, solution optimization, and evaluation. Details are shown in Figure 9.2.

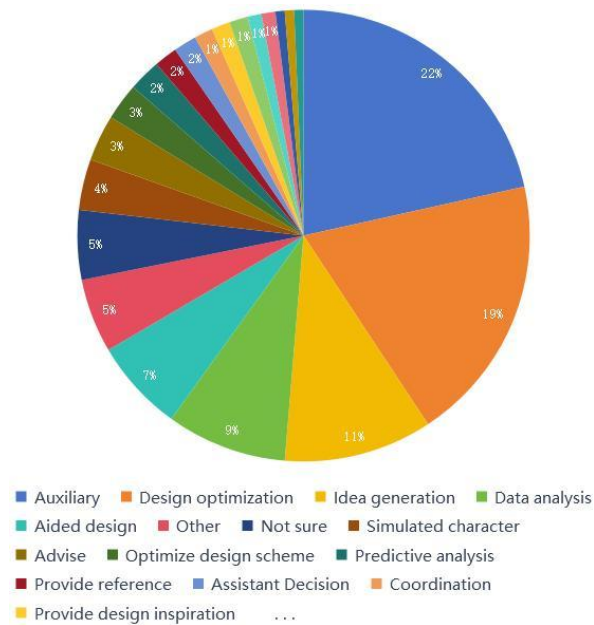


Figure 9.2 The Role of AI in Conceptual Design Evaluation

Designers are positive about the application of AI, but at the same time, they are concerned about its accuracy and credibility. Figure 9.3 shows that designers believe the most important AI design evaluation functions include data analysis, creative generation, and solution optimization.

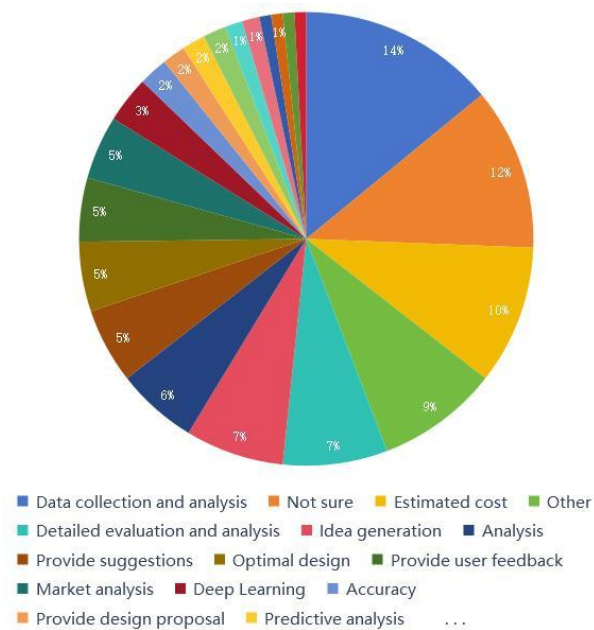


Figure 9.3 AI Concept Design Evaluation Expected Functions

The study found that designers believe improving the accuracy of AI algorithms, increasing data training, and optimizing algorithms is necessary. The aspects that need to be improved in the AI concept design evaluation model are shown in Figure 9.4. Designers are not satisfied with the accuracy and credibility of AI's current algorithms. 22% of designers mentioned that AI needs to receive more data training and provide more data. 21% of designers believe that AI needs algorithm optimization. In addition, improving the accuracy of evaluation and simplifying the operation process are also frequently mentioned needs.

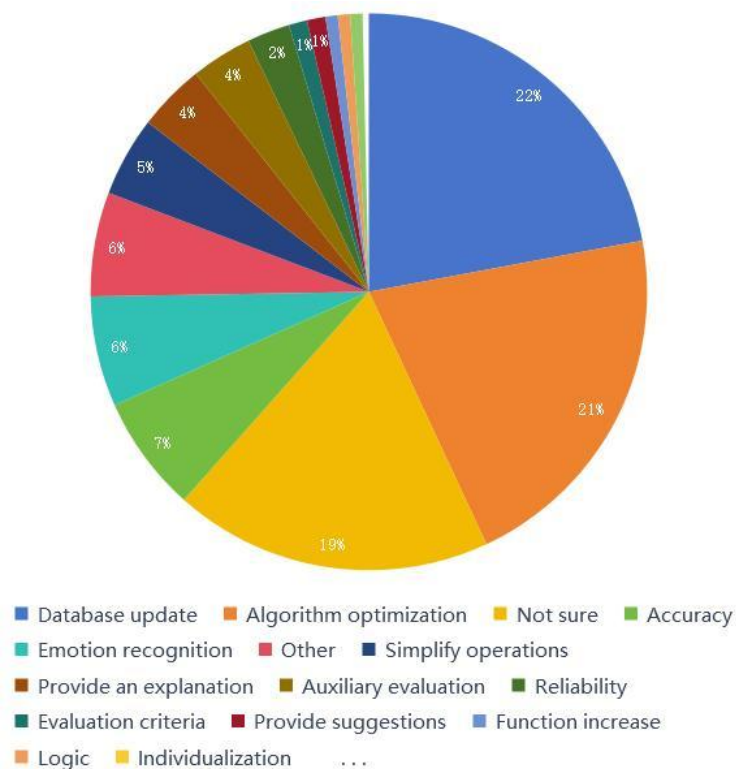


Figure 9.4 Needs of AI Concept Design Evaluation

Figure 9.5 shows the designers' suggestions for improving the application of AI in design evaluation. They suggested strengthening algorithm training, improving system transparency, and providing evaluation suggestions in

specific scenarios. The most popular suggestion was to strengthen the training and optimization of algorithms to improve the accuracy of AI evaluation tools and to update the database and iterate the algorithm in real-time. Regarding usage, designers hoped there would be a detailed explanation of the working principle and operation tutorials. Designers want to get design evaluations and suggestions in specific scenarios and specific design tasks rather than empty and unrealizable analysis results and suggestions.

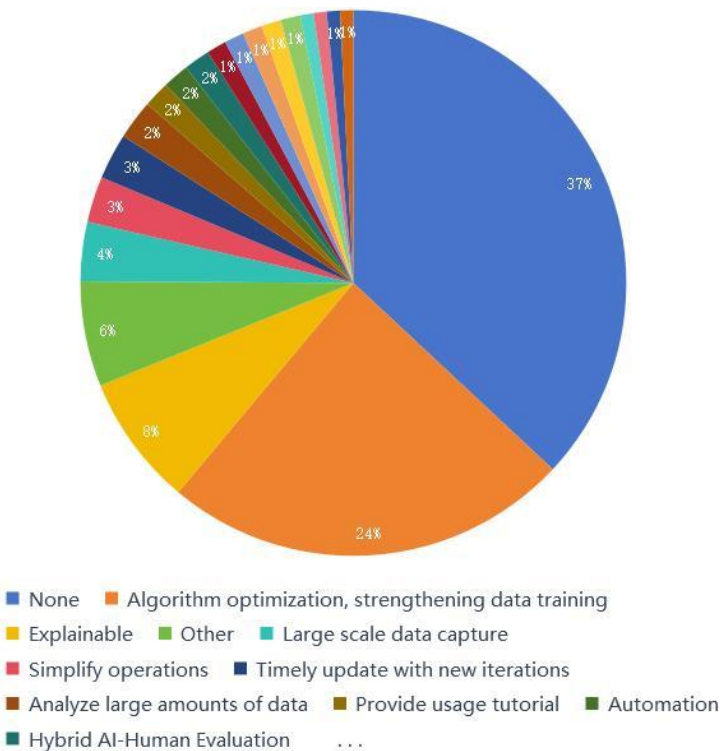


Figure 9.5 Suggestions for Improving AI Design Evaluation

Regarding the expectations and suggestions for the future development of AI in design evaluation, designers expect AI to improve work efficiency and design quality and provide more objective and feasible suggestions. Designers have expectations for AI. Specifically, they hope AI can improve work efficiency, design quality, and have models for each field. Moreover, they expect AI to generate more objective and feasible suggestions. The

details are shown in Figure 9.6. However, users of open-ended questions filled out None, which indirectly reflects that users are unfamiliar with AI products. This result also shows that there is still room for research in AI-assisted design.

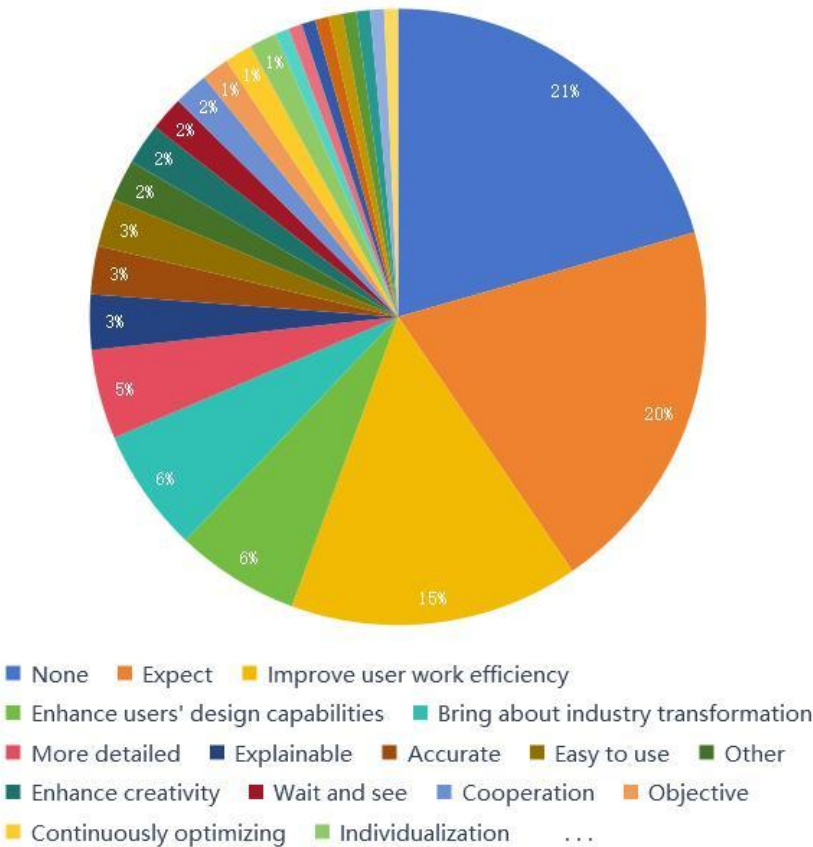


Figure 9.6 Expectations and Suggestions for AI in Design Evaluation

The coding results of open questions show that designers believe that AI should support conceptual design evaluation. Designers are positive about the application of AI but at the same time, have concerns about its accuracy and credibility. To improve the application of AI in design evaluation, designers suggest strengthening algorithm training, improving system transparency, and providing evaluation suggestions in specific scenarios. For the future development of AI, designers expect it to improve work efficiency and design quality and provide more objective and feasible suggestions. It is

worth noting that the study also found that the Relationship between designers' familiarity with AI and their confidence and willingness to use it is not significant. This finding differs from previous research results and emphasizes the need for more comprehensive strategies to promote the application and acceptance of AI tools in design.

In summary, the questionnaire survey shows that designers have concerns about the application of AI. These concerns include data quality, lack of trust in AI, and lack of industry standards, indicating that designers face multiple challenges when using AI design evaluation tools. The 129 Chinese designers have a cheerful outlook toward applying AI in concept design evaluation, but there are significant differences in their trust in accuracy. Trust is a key factor affecting designers' willingness to use AI evaluation tools. Improving designers' trust in AI tools can effectively promote the application and popularization of these tools. The Relationship between designers' familiarity with AI and their confidence and willingness to use AI is insignificant, indicating that familiarity is not the main factor affecting their confidence and willingness to use it. The Relationship between designers' work experience in the design industry and their confidence and willingness to use AI could be more robust.

9.3.2 Interview Coding Results

Three researchers conducted a three-level coding analysis of grounded theory based on the interview data, including open coding, axial coding, and selective coding. This coding method allows researchers to extract key concepts from the original data and organize these concepts at a higher level

of abstraction. The following are the specific implementation and results of the coding.

Through open coding, this study identified initial concepts such as AI as a design student assistance tool and the limitations of AI assessment. The details are shown in Table 9.12.

Table 9.12 Open Codes of Interview

Open Codes	Original sentence examples (e.g.)
AI as design students' assistance tool	As a student, you may not know whether your design work is good or bad during your student period. So, it is especially useful to have such an AI-assisted evaluation tool. AI can evaluate the design work.
Limitations of AI assessment	Sometimes a product is not just an appearance, but also has functions or interactive aspects. AI cannot accurately identify and feel it like a human by relying on a static picture, right?
Human-computer collaborative assessment	If there is an evaluation from an expert's perspective, plus an evaluation from AI, you may get a comprehensive result. It is more accurate than simply evaluating the design by humans.
Trust in AI assessment	I think we should give users credible data. For example, in addition to the rating results, there are rating reasons or criteria.
Directions for improvement of AI assessment	This kind of products that purely display beauty, such as jewelry, cannot be evaluated together with electronic products. Because their evaluation standards are different, these products must be classified in advance.

In the axial coding stage, the researchers organized these concepts into broader categories, such as, AI's auxiliary role and Human-machine collaboration mode, as shown in Table 9.13.

Table 9.13 Axial Codes of Interview

Axis codes	Related Open codes
AI's auxiliary role	AI as design students' assistance tool, Limitations of AI assessment
Human-machine collaboration mode	Human-computer collaborative assessment, Trust in AI assessment
AI model optimization	Limitations of AI assessment, Directions for improvement of AI assessment

Finally, in the selective coding stage, the researchers identified core themes such as, the positioning of AI in design evaluation and the challenges and developments of AI evaluation. Table 9.14 is the result of selective coding.

Table 9.14 Selective Codes of Interview

Selective Codes	Related Axis codes
The positioning of AI in design evaluation	AI's auxiliary role, Human-machine collaboration mode
The challenges and developments of AI evaluation	Human-machine collaboration mode AI model optimization

This analytical approach revealed the mixed views of respondents on AI-assisted design evaluation. On the one hand, participants recognized the potential of AI as an auxiliary tool, especially in helping students evaluate their design works. On the other hand, they also expressed concerns about the evaluation capabilities of AI, especially in understanding the function and emotion of the design. The human-computer collaboration model was seen as an ideal evaluation method, which suggests that future AI-assisted design evaluation systems should focus on human-computer complementarity rather

than rely solely on AI. In addition, improving the credibility and accuracy of AI evaluation is considered a key direction for future development.

These findings provide important insights for developing more effective AI-assisted design evaluation tools while highlighting the importance of maintaining human involvement in AI evaluation. Future research can further explore optimizing the human-computer collaboration model and improving AI's ability to understand complex design concepts.

9.4 Summary and Discussions

Through quantitative and qualitative interviews, this study used a mixed methods approach to explore designers' attitudes toward, trust in, and willingness to use artificial intelligence-assisted design evaluation tools. The results revealed key insights for applying artificial intelligence in the design evaluation.

In design evaluation and validation, the integration of artificial intelligence is becoming increasingly common, providing a range of advantages in various domains. The study by Niraula et al. (2024) explored the dynamics of human-AI interaction in the decision-making process, emphasizing the positive correlation between the level of agreement with AI recommendations and trust in AI. This correlation highlights the importance of building trust and maintaining consistency with artificial intelligence systems to use their recommendations effectively. Artificial intelligence assistants are increasingly integrated into design to support the creative process and improve productivity. Peuter et al. (2021) emphasized the importance of AI assistants

that align with the goals and abilities of designers, as these assistants can foster symbiotic relationships and promote innovation.

The collaboration between artificial intelligence assistants and human evaluators is an emerging research field, as demonstrated by studies such as Kuang et al. (2023), Lankes and Stockl (2023). These studies explore how artificial intelligence can assist in user experience (UX) evaluation and game design processes, thereby gaining a deeper understanding of the potential of AI in simplifying evaluation tasks and enhancing design outcomes. By utilizing the artificial intelligence capabilities in these fields, practitioners can benefit from enhanced analysis tools and decision support systems to optimize the design process.

In short, integrating artificial intelligence into design evaluation, decision-making, and collaborative settings has enormous potential for improving efficiency, accuracy, and creativity in different fields. By utilizing artificial intelligence capabilities to enhance human expertise and decision-making processes, practitioners can unlock new opportunities for innovation and optimization in their respective fields. This study also focuses on this popular research field, hoping to explore how artificial intelligence empowers product design evaluation to optimize the design process.

This study used a mixed methods approach combining quantitative and qualitative interviews to explore designers' attitudes toward, trust in, and willingness to use AI-assisted design evaluation tools. The results revealed key insights for applying artificial intelligence in the design evaluation.

First, the quantitative analysis showed designers' cheerful outlook toward using AI in conceptual design evaluation. This is consistent with previous research highlighting the potential benefits of AI in enhancing the design process. However, the study also found significant individual differences in trust, especially in terms of trust in the accuracy of AI evaluation. This difference in trust suggests that targeted approaches are needed to build trust for different user groups.

The study found no significant correlation between designers' familiarity with AI and their confidence or willingness to use AI tools. This finding challenges the common assumption that increased familiarity leads to higher acceptance and highlights the complexity of factors affecting the adoption of AI in design. Similarly, there needed to be a stronger correlation between work experience and AI acceptance. This result suggests that other factors are more critical in driving adoption, such as the perceived usefulness and ease of use of AI systems.

Qualitative interviews provide deeper insights into designers' perceptions of the role of AI in design evaluation. The results cross-validated that designers primarily view AI as an auxiliary tool, especially in data analysis, idea generation, and solution optimization. This is consistent with human-machine collaboration, where AI enhances rather than replaces human creativity and decision-making. However, designers also expressed concerns about AI's ability to understand complex design concepts, especially those related to functionality and emotion. This highlights the need to develop AI systems to explain and evaluate subjective design elements better.

A key finding from quantitative and qualitative data is the importance of transparency and explainability in AI systems. Designers stressed the need for clear explanations of how AI works and the criteria for evaluation. This is consistent with recent research on explainable AI and its role in building trust and promoting human-machine collaboration. Based on these findings, this study can make the following recommendations for improving AI-assisted design evaluation tools.

- Improving the transparency and explainability of AI systems to build trust and promote better understanding among designers.
- Develop more sophisticated AI models to evaluate the design's subjective and emotional aspects through advanced natural language processing and computer vision technologies.
- Focusing on creating human-machine collaborative systems that leverage the strengths of human designers and AI rather than pursuing fully autonomous AI evaluations.
- Providing designers with tailored training and support to effectively integrate AI tools into their workflows and address individual differences in AI acceptance.

This study has limitations. The sample was limited to Chinese designers, which may affect the applicability of the findings to other cultural contexts. In

addition, the study focused on current cognition and did not longitudinally track changes in attitudes toward AI over time. Future research could explore cross-cultural differences in designers' acceptance of AI, investigate the long-term impact of AI adoption on design practice, and develop and test AI systems that can address the concerns and suggestions raised by designers in this study. In addition, exploring the ethical implications of AI in design evaluation, especially regarding data privacy and potential bias in AI systems, would be a valuable direction for future research.

In summary, Study 5 further explored users' experience and satisfaction with AI-assisted design evaluation tools through a mixed method (questionnaire and structured interview). This study not only confirmed the findings of the earlier chapters, especially the experimental results of Study 4 but also offered valuable insights into the practical application of AI tools. Although the study may have sample limitations, its results offer essential guidance for future improvements and research directions, laying the foundation for future discussions and summaries. The next chapter will conduct a comprehensive discussion and reflection based on Study 1- Study 5 findings.

Chapter 10. Discussion

This chapter discusses the findings from the entire research process, with a particular focus on the limitations of the research and future research directions. The thesis reflects on the possible applicability limitations of research based on concept design data. It explores expanding the research to a broader range of industry scenarios. This discussion summarizes the earlier chapters' research results and provides constructive suggestions for future AI-assisted product design evaluation research.

10.1 General Research Results

In this thesis, a series of studies systematically explored the potential application of artificial intelligence (AI) in product design evaluation and its actual effects. With the rapid development of AI technology, its application has begun to change traditional processes and methods. Especially in the field of product design, the introduction of AI is expected to not only improve design efficiency but also improve design quality through new data analysis methods. This study focuses on the issue of how AI can assist product design evaluation, aiming to fill the gap in the existing literature and provide new perspectives for design practice.

The study's main objective is to verify the effectiveness of the AI-assisted design evaluation system (AiDE) and explore its application in actual design processes. This study examines the model's performance on multimodal datasets by building and verifying a deep learning-based evaluation model. It

compares the performance of human experts and AI in design evaluation tasks through human-machine comparison experiments. In addition, the study also covers the application potential of AI in design education, as well as designers' acceptance and trust in AI tools.

This chapter aims to explore the research results in depth, analyze their theoretical significance and practical applications, and discuss the study's limitations. Based on the research findings, this chapter will discuss the practical utility of AI in product design evaluation, evaluate its application prospects in education and industrial practice, and propose directions for future research. Through these discussions, this study aims to provide valuable insights and suggestions for researchers and practitioners in the design field. This section begins with the discussion of answering RQs1-4.

RQ1: At which product design stage does AI need auxiliary tools the most? What functions should these tools have?

Through the implementation of the AiDE system and in-depth user feedback analysis, this study verified that AI tools are particularly useful in the concept and prototype stages of product design. AI assistance can help designers quickly evaluate and iterate design ideas, optimize design solutions, and enhance designs' innovation and market adaptability in the initial design concept stage. AI tools provide functions such as data analysis, idea generation, and solution optimization, consistent with the expectations of research question RQ1.

RQ2: What data types need to be collected and processed to develop AI-assisted design evaluation tools?

The results show that building an efficient AI evaluation model requires multimodal data, including images and text. Image data helps the model understand the visual features of the design. In contrast, text data (such as design instructions and user feedback) provides in-depth insights into design intent and functionality. In addition, through human-computer comparison experiments, this study found that the quality and processing of data directly affect the accuracy and reliability of AI evaluation, which answers RQ2.

RQ3: How to build an AI product design evaluation model based on multimodal data to improve the accuracy and comprehensiveness of the evaluation?

The AI model developed in this study uses deep learning technology to process and fuse multimodal data. The model can more comprehensively understand and evaluate design solutions by combining image and text analysis. The experimental results show that the combined multimodal model performs better than the single modality model in most evaluation dimensions, especially in technology and market evaluation. This proves that the research direction of RQ3 is correct.

RQ4: How effective is the AI evaluation model? In practical applications, what are the similarities and differences between the model and the evaluation results of human experts?

Through human-machine comparison experiments, this study evaluated the performance of AI models and human experts in design evaluation. The results show that the AI model is consistent with human evaluation in technical and market evaluation, but there is still a gap in design innovation and aesthetic evaluation. These findings point out the potential and

limitations of AI evaluation tools in practical applications, providing detailed answers to RQ4.

This section will provide an in-depth discussion of the advantages of AI models and how they can collaborate with human designers based on these research findings.

10.1.1 Advantages of AI Models

This study aims to develop and validate an AI-assisted product design evaluation system through in-depth exploratory studies. This section will detail the advantages of AI models in improving design efficiency, enhancing innovation capabilities, and managing complex evaluation tasks in combination with references and research results. These advantages reflect the transformative potential of AI technology in product design and provide important directions for the development of future design methods and tools.

- **Improving design efficiency**

AI technology shows enormous potential to improve design efficiency. As Quan et al. (2023) pointed out, integrating big data and AI algorithms revolutionizes product design methods, enabling designers to process and analyze massive amounts of information to optimize design outcomes. The experimental results of this study confirm this view. For example, the experiments in Chapter 7 show that the accuracy of the ResNet18+BERT model in the technical evaluation dimension reaches 73.01%, significantly improving the evaluation efficiency. This efficient data processing

capability enables designers to make decisions more quickly and shorten the product development cycle.

In addition, the multimodal product design dataset constructed in Chapter 6 demonstrates AI's ability to integrate and analyze data. This echoes the research of Zhou et al. (2020), who emphasized the importance of multimodal data fusion in improving the overall performance of AI systems. By integrating multiple forms of design data, such as images and text, AI systems can more comprehensively understand and evaluate product designs, thereby accelerating the design evaluation process.

- Enhancing design innovation capabilities

AI improves design efficiency and shows enormous potential in enhancing design innovation. The survey results in Chapter 5 show that 72.05% of designers have used AI creative tools, which shows that AI has become useful for promoting design innovation. This finding is consistent with the research of Verganti et al. (2020), who explored how AI can fundamentally change the innovation and design process by improving scalability, expanding boundaries, and enhancing adaptability.

In this study, AI models demonstrated unique innovation-promoting capabilities. For example, the multimodal AI evaluation model developed in Chapter 7 can simultaneously analyze a product's visual and descriptive features to provide a more comprehensive evaluation result. This capability enables AI to capture subtle innovations that human designers may overlook, thereby inspiring modern design ideas. In addition, by analyzing a large amount of historical design data, AI models

can provide designers with novel creative inspiration, which echoes the AI-driven product design framework proposed by Liu and Kim (2023).

- **Managing complex assessment tasks**

AI performs well in handling complex design evaluation tasks, especially in processing multi-dimensional, unstructured data. The experimental results in Chapter 7 show that multimodal models that combine image and text information (such as ResNet18+BERT) perform best in the market evaluation dimension (ACC=0.723, SRCC=0.5874, PLCC=0.5718). This finding is consistent with the research of Radu et al. (2018), who emphasized the advantages of multimodal deep learning in handling complex data sets and reasoning tasks.

The AI model can simultaneously consider multiple aspects, such as functionality and aesthetics, to provide comprehensive evaluation results. This ability is particularly important when dealing with highly subjective design dimensions. For example, the human-machine comparison experiment in Chapter 8 showed that in evaluating the design dimension, the accuracy of the optimized AI model reached 82.35%, much higher than the 37.73% of the initial model. This shows that AI has the potential to provide valuable insights into extraordinarily complex and subjective evaluation tasks.

10.1.2 Collaboration between AI and Human Designers

Although AI has shown significant advantages in product design evaluation, the results show that a human-machine collaborative model is the ideal evaluation method. This is consistent with the findings of Chong et al. (2022),

who emphasized that human designers' effective use of AI input is critical to success.

- Complementary advantages

The results show that AI models excel in processing objective data and large-scale information, while human designers have creativity and subjective judgment advantages. The interview results in Chapter 9 show that designers believe the ideal evaluation model should be human-computer collaboration, with AI as an auxiliary tool to provide data support and preliminary evaluation. This view echoes the research of Peuter et al. (2021), who emphasized that AI assistants consistent with designers' goals and abilities can promote symbiotic relationships and drive innovation.

- Improving design quality

Through human-machine collaboration, the data analysis capabilities of AI and human creativity can be combined to improve design quality. The experimental results in Chapter 8 show that in the evaluation of design dimensions, the accuracy of the optimized model (82.35%) is significantly higher than that of AI or manual evaluation alone. This finding is consistent with the evaluation model based on fuzzy dual experience proposed by Chen (2016), which emphasizes the importance of combining engineering design with customer response.

- Enhancing assessment accuracy

AI models can provide objective data analysis and preliminary evaluation for human designers, while human designers can interpret and adjust AI results based on experience. The questionnaire survey in Chapter 9 showed that 61.24% of designers believed AI could effectively assist design decision-making. This collaboration improves the accuracy of the evaluation. It enhances the interpretability of the evaluation results, consistent with the importance of explainable AI to human decision-making emphasized by Alufaisan et al. (2020).

- Optimizing workflow

AI can take on repetitive data processing and preliminary evaluation work, allowing designers to focus on more creative tasks. The survey results in Chapter 5 show that 55.22% of designers hope that AI will assist the design process as an interactive tool. This division of labor can improve overall design efficiency while allowing human designers to improve their creativity. This echoes the human-computer interactive collaboration framework in data-driven text generation proposed by Strobelt et al. (2022).

- Promoting learning and development

Designers can continuously learn new evaluation methods and design trends by collaborating with AI. In the interviews in Chapter 9, designers expressed the hope that AI could provide design advice and market trend analysis to help them continuously improve their professional capabilities. This model of continuous learning and improvement is of great significance to the long-term career development of designers and is

consistent with the human-computer collaborative learning model proposed by Schoonderwoerd et al. (2022).

The results show that AI has significant advantages in product design evaluation, especially in improving efficiency, enhancing innovation, and managing complex evaluation tasks. However, designers believe that AI cannot completely replace human designers; instead, it should be used as a powerful auxiliary tool. Through human-machine collaboration, the data processing capabilities of AI and human creativity can be fully utilized, thus achieving a higher quality and more efficient product design evaluation process.

10.2 Discussion of the Challenges of AI in Subjective Evaluation Tasks and the Extension of Future Research

Based on the research results of the previous nine chapters, this section will discuss the application effects and potential challenges of artificial intelligence (AI) in product design evaluation, especially the effectiveness of multimodal data fusion and machine learning techniques.

10.2.1 Application Effect of AI in Product Design Evaluation

According to the research results in the previous chapters, AI technology has shown significant potential in product design evaluation. Through automated data analysis and pattern recognition, AI can quickly identify key elements in the design, thereby providing immediate feedback and suggestions. For example, in Chapter 8, the AI-assisted evaluation system (AiDE) can use deep

learning models to process design images and text descriptions to provide a comprehensive design evaluation.

However, the application of AI in product design evaluation also faces challenges. The research in Chapter 8 shows that although AI performs well in processing technical and market evaluations, its performance is sometimes limited when dealing with design tasks involving highly creative and subjective aesthetic judgments. This is due to the subjective nature of design evaluation, and it is difficult for AI models to fully understand and simulate the creative process and aesthetic preferences of human designers.

10.2.2 Effectiveness of Multimodal Data Fusion

Multimodal data fusion refers to the comprehensive use of different forms of data, such as images and text, in AI systems. Chapters 6 and 7 show that by fusing image and text data, AI evaluation models can more fully understand design intent and context, thereby improving the accuracy and depth of evaluation.

For example, by analyzing design drawings (image data) and design instructions (text data), AI can comprehensively consider the visual beauty and functionality of the product and provide a more comprehensive and objective evaluation. This multimodal fusion enhances the AI model's explanatory power and makes the evaluation process more transparent and credible.

10.2.3 The Role of Deep Learning Technology in Improving Design

Evaluation

Machine learning technology is the core of implementing AI evaluation models. The experiments in Chapter 7 show that using deep learning algorithms, such as multimodal fusion of computer vision and natural language processing technology, can significantly improve the efficiency and quality of design evaluation.

By training AI models to identify and learn design patterns and trends, machine learning technology can automatically perform tedious analysis tasks, freeing designers from repetitive labor and focusing on more creative design work. In addition, AI's continuous learning and adaptability mean that it can continuously improve the accuracy and relevance of evaluations over time.

10.2.4 Challenges and Future Research Directions

This study aims to develop and validate an AI-assisted product design evaluation system through in-depth exploratory studies. Although AI has shown significant potential in product design evaluation, challenges still exist. Significant challenges are still faced when performing highly subjective evaluation tasks, especially in evaluating design dimensions. These challenges involve model transparency and interpretability, data quality, and diversity. The following will combine the findings of this study with relevant literature to explore these challenges and future research directions in depth.

10.2.4.1 Transparency and Interpretability of Model

The transparency of AI models in design evaluation is a key factor affecting their widespread acceptance and trust. As shown in the experimental results of Chapter 8, in the evaluation of the design dimension, the performance of the AI model is weak (ACC is 0.3773), reflecting the challenges AI faces in dealing with highly subjective evaluation tasks. Design evaluation often involves complex subjective judgments, such as judging whether a design is beautiful and whether it meets user needs or market trends. However, the decision-making process of AI models is often a black-box operation, lacking sufficient transparency and explainability, which may cause designers and stakeholders to be skeptical of the evaluation results provided by AI.

This finding is consistent with the findings of Herskovits (2021), who noted that AI models have difficulties in interpretability, which may cause users to doubt their reliability and validity. Similarly, Khanolkar et al. (2021) also emphasized that although AI can assist engineering design, a fully automated design process remains elusive, highlighting the importance of maintaining a human-machine collaborative approach.

Future research could explore developing more intuitive visual explanation tools to improve the transparency and interpretability of AI models. These tools can clearly show how AI extracts feature from design data and makes evaluation decisions based on these features. This suggestion echoes the idea of interpretable visual aesthetics datasets proposed by Chen et al. (2020), also emphasized that although AI can assist engineering design, a fully automated design process still needs to be discovered, highlighting the importance of maintaining a human-machine collaborative approach.

10.2.4.2 Quality and Diversity of Data

Data is the cornerstone of AI model performance, and the quality and diversity of the dataset directly affect the accuracy of AI evaluation. In this study, Chapter 6 details the construction process of the multimodal product design dataset. However, as shown in the questionnaire survey results in Chapter 9, 61.24% of designers believe that insufficient data quality is one of the main problems facing current AI-assisted design evaluation. This finding echoes the research of Zhang et al. (2022), who emphasized the importance of understanding consumers' reactions to AI-designed products in developing effective evaluation models.

In product design evaluation, especially when dealing with design innovation and user experience, high-quality and diverse data is needed to capture a wide range of user preferences and market trends. However, existing datasets are often biased or fail to fully cover all potential user groups, which may result in AI models failing to reflect the needs and preferences of different users accurately. This issue is also reflected in the study of Quan et al. (2023), who pointed out that the fusion of big data and AI algorithms is revolutionizing product design methods, enabling designers to process and analyze massive amounts of information to optimize design results.

10.2.4.3 Discussion of Future Research Directions

In response to the challenges of AI in subjective evaluation tasks, future research can explore more advanced AI algorithms to improve the ability to manage complex and subjective design tasks. At the same time, developing more effective multimodal data integration methods and improving the

transparency and user trust of AI systems will be the key to promoting AI's application in product design.

- Enhancing model transparency and explainability

Develop new visualization tools and explanatory algorithms to help users better understand AI's evaluation process and basis. This direction echoes the research that explored the impact of explainable AI on human decision-making (Alufaisan et al., 2020).

- Improving data quality and diversity

Future research could expand the scope of data collection to ensure that the dataset can fully reflect diversified product types and market changes.

- Introducing marketing data and large-scale user reviews

In the future, AI evaluation's market sensitivity and accuracy can be enhanced by analyzing actual sales data or developing user online evaluation functions to provide real-time feedback.

- Optimizing the interaction of artificial intelligence models

Future interactive designs can provide immediate visual or sound feedback to help users understand AI operations and responses. For example, from human-computer interaction to deep coordination, trust building is an important breakthrough (Song & Luximon, 2020).

- Exploring more application scenarios and more advanced AI algorithms

Applying AI to more design and creative work areas is another important direction for future research. By exploring the application of AI in these new areas, future research can not only expand the scope of the application of AI technology but also further promote design innovation and technological development. Future research can explore more advanced AI algorithms, such as reinforcement learning, to improve AI's ability to manage complex and subjective design tasks. Applying these algorithms will help improve the performance of AI systems in product design evaluation, enhancing their creativity and adaptability.

Through these research directions, future AI models will be able to perform design evaluation tasks more accurately and provide more valuable insights into design innovation and user experience. This will help better support designers and companies in standing out in the fiercely competitive market while promoting AI's in-depth application and development in product design.

Although AI has shown enormous potential in product design evaluation, it still faces major challenges in dealing with highly subjective evaluation tasks. By improving the transparency and interpretability of models, improving data quality and diversity, and exploring new research directions, future AI-assisted design evaluation systems can better meet the needs of designers, promote design innovation, and bring users a better product experience.

10.3 Discussion of Exploring and Optimizing the Human-computer Collaboration Model

This section will explore in depth how to optimize the human-computer collaboration model, especially improving the human-computer interface and enhancing the credibility and explainability of AI-assisted evaluation.

10.3.1 Improving the Human-computer Interface

The GUI is the direct medium for designers to interact with AI tools, and its design quality directly affects the practicality and effectiveness of AI tools. Based on the experimental results in Chapter 8, this study found that when designers use AI tools, intuitive and easy-to-use interfaces can significantly improve work efficiency and reduce operational errors. The research results in Chapter 9 show that designers hope to achieve a clearer key link. The following are key optimization directions.

- **Complying with the designer's workflow**

The human-computer interface should be designed to fit naturally into the designer's daily workflow. For example, the interface layout can mimic trendy design software, using familiar toolbars, menus, and shortcut keys, allowing designers to seamlessly switch between AI tools. This design concept echoes the view proposed by Peuter et al. (2023) that AI assistants should be aligned with the designer's goals and abilities.

- **Customization and personalization**

Providing customized interface setting options allows designers to adjust the interface layout and functional modules according to personal preferences and specific project needs. For example, allowing users to adjust the evaluation criteria according to project types, such as industrial or fashion design. This personalized setting not only improves the user experience but also enhances the designer's sense of control over the AI tool, which is consistent with the research findings on human-computer collaboration in UX evaluation (Kuang et al., 2023)

- Interactive feedback and real-time updates

The interface should be able to display the progress and results of AI analysis and evaluation in real-time, help designers instantly understand the working status of AI tools, and provide detailed improvement reports. This real-time interaction and feedback mechanism is consistent with the comprehensive framework for enhancing the performance of human-machine collaboration (Jiang et al., 2022).

10.3.2 Enhancing the Credibility and Interpretability of AI-assisted Evaluation

The effectiveness of AI in design evaluation relies heavily on designers' trust in its output. Therefore, improving the transparency and explainability of AI systems is key. This view echoes research on improving the trustworthiness of AI solutions (Vianello et al., 2023). The following are key directions for improvement.

- Improving transparency

AI systems should clearly explain their decision logic and their data sources. For example, when an AI tool makes a design suggestion, it should explain how the suggestion was extracted from user data. This transparency will enhance designers' trust in AI systems and help them better understand and use AI outputs.

- Adding visual explanation tools

Visualization tools are developed to depict how AI processes input data and draw conclusions. For example, the scoring criteria can be communicated to designers as a text report using natural language generation tools.

- User education and training

By providing regular training and education programs, designers can better understand the principles and potential of AI technology. Through seminars, online courses, and interactive tutorials, designers can be educated on effectively using AI tools for design evaluation. This educational approach is consistent with discussion on integrating emerging technologies into future designer education and helps to cultivate designers' AI literacy and human-computer collaboration capabilities (McCardle, 2002).

Through the above measures, designers can improve their proficiency and comfort in using AI tools and enhance their trust and acceptance of AI

evaluation results. These improvements will help achieve more effective human-computer collaboration and promote innovation and efficiency in design practice. Future research will continue to explore new human-computer interaction technologies and methods to continuously optimize the human-computer collaboration model and better serve the needs of the design industry.

In short, optimizing the human-computer collaboration model is an ongoing process that requires simultaneous advancement in multiple aspects, such as technological innovation, user experience design, and education and training. By continuously improving the human-computer interface, enhancing the credibility and explainability of AI systems, and strengthening designers' AI literacy, future research can build a more efficient and innovative human-computer collaborative design ecosystem and promote the continued development and progress in product design evaluation.

10.4 Limitations and Future Research Directions

In this study, the thesis systematically explores the application of artificial intelligence technology in product design evaluation. It evaluates the effectiveness and practicality of a multimodal AI evaluation system (AiDE) by building and evaluating it. However, despite the achievements of the study, there are still limitations that provide new directions for future research.

10.4.1 Limitations

The thesis explores the application of AI in product design evaluation. It demonstrates AI's potential value and effectiveness in the design process by

building and experimenting with a multimodal AI evaluation system. However, although this study provides useful insights, it also has inevitable limitations. These limitations are reflected in the scope of sample selection, research methods' focus, and AI models' transparency and interpretability. Identifying these limitations is important for correctly interpreting research results, avoiding over-generalizing research conclusions, and guiding future research directions.

- Insufficient sample diversity

This study relies on a specific sample set, including student design works and professional designers' evaluations. The number and diversity of samples are limited, which may affect the broad applicability and generalizability of the results. Future research needs to verify the effectiveness of AI evaluation models in a wider and more diverse range of design fields.

- Limitations of methods

Although multimodal data and deep learning techniques are used, the research focuses on integrating image and text data. It insufficiently explores other potentially useful data modes (such, sound, video, etc.). In addition, the research relies on quantitative evaluation methods, and the treatment of subjective and emotional factors in design evaluation may not be in-depth enough.

- AI models lack transparency and explainability

Although current AI evaluation models perform well in evaluating indicators such as technology and market, their decision-making process lacks transparency and explainability, which may affect users' trust and acceptance of AI evaluation results.

Overall, this study has made progress in exploring the application of AI in product design evaluation. However, due to factors such as sample selection restrictions, research method limitations, and lack of transparency of AI models, its conclusions' general applicability and deep explanatory power have been affected to a certain extent. Future research needs to expand the scope and diversity of samples, adopt more comprehensive research methods, and focus on improving the interpretability of AI models to improve the effectiveness and practicality of the research. Through these efforts, AI technology can help to optimize product design while providing designers and researchers with more accurate and reliable evaluation tools.

10.4.2 Future Research Directions

While AI has shown exciting potential in product design evaluation, this study suggests that AI needs to be more deeply integrated into the design process by improving the quality of evaluations. Future research must address issues that have yet to be fully explored in existing research and develop new methods to improve the performance and transparency of AI systems. This will include a more extensive exploration of the role of AI in design evaluation, improving the diversity and quality of data, enhancing the interpretability of models, and extending AI evaluation models to a wider range of design domains.

- Improve the breadth and diversity of samples

Future research can apply AI evaluation models to more types of design projects, such as industrial design, fashion design, interface design, etc., to test and verify the versatility and effectiveness of the model. In addition, increasing the diversity of samples, such as works with distinct cultural backgrounds and design styles, can improve the comprehensiveness and applicability of the research.

- Explore new data models and research methods

Future research can consider integrating more data models, such as videos, 3D models, and real-time feedback data, which may provide richer contextual information for product design. At the same time, exploring qualitative research methods, such as in-depth interviews and case studies, can help better understand how designers and users feel and think about AI evaluation tools.

- Expand the interpretability of AI evaluation models

Research should develop new algorithms and techniques to improve the transparency and interpretability of AI evaluation models. For example, using interpretable machine learning techniques (such as decision trees and rule engines) to enhance the interpretability of models or developing visualization tools to reveal how AI makes evaluation decisions.

- Interdisciplinary collaborative research

Artificial intelligence and product design are both interdisciplinary subjects. Future research can collaborate with other disciplines, such as psychology, sociology, and art, to explore how AI can better serve creative generation, emotional expression, and social interaction.

In summary, this study has opened new perspectives on the application of AI in product design evaluation and pointed out directions that can be further explored in future research. Future research should focus on increasing the diversity of data sets, improving the transparency and explainability of AI models in the evaluation process, and exploring the applicability of AI technology in distinctive design fields. In addition, interdisciplinary collaboration will be key to promoting innovative applications of AI in design evaluation. Through these efforts, AI technology will better serve the needs of designers and promote the optimization of design processes while ensuring that technological advances can be applied responsibly at ethical and social levels.

This chapter comprehensively discusses the findings from the entire research process, with a special focus on the limitations of the research and future research directions. By reflecting on the possible applicability limitations of research based on competition data, this study suggests expanding the research scenario to a wider range of industry scenarios. This discussion summarizes the research results from the earlier chapters. It provides constructive suggestions for future AI-assisted product design evaluation research, providing an important basis for writing the conclusion chapter. The next chapter will draw a conclusion based on these discussions.

Chapter 11. Conclusion

This concluding chapter summarizes the main findings of the entire study and discusses the theoretical and practical contributions. The thesis proposes a model for AI-enabled product design concept evaluation, which is the core contribution of this study. By reviewing five interrelated studies (Study 1-5), the thesis shows how AI can effectively support and enhance the product design evaluation process, providing new perspectives and directions for future developments in the design field.

11.1 Main Research Findings

This thesis systematically explores the application of AI in product design evaluation, significantly enhancing the efficiency and quality of design evaluation through multimodal data and deep learning technology. This study designed a hybrid research method, combining quantitative and qualitative analysis, to comprehensively evaluate the performance of AI evaluation models and their application effects in the actual design process. The study deeply analyzed the practical application of AI in the design process through five phases of research (Study 1-Study 5). It proposed an AI-assisted product design evaluation system based on multimodal data. This study demonstrated the potential of AI in design evaluation and challenged and improved the existing design evaluation methods.

In the field of product design, the application of AI can improve design efficiency and innovation capabilities. (2023) pointed out that combining big

data and AI algorithms revolutionizes product design methods, enabling designers to process and analyze copious amounts of information and optimize design results. Based on this theory, this paper further explores the specific application of AI in design evaluation, primarily how to process and analyze multimodal data to provide more comprehensive evaluation results.

11.1.1 Development and Application of Multimodal AI Evaluation Models

In the current design field, with the rapid development of technology, especially the advancement of artificial intelligence technology, increased research has begun to explore how to apply AI technology to product design evaluation. This study successfully developed a new multimodal AI evaluation model that can integrate visual and text data to comprehensively capture and deeply evaluate all aspects of product design. Through the analysis of visual images and the interpretation of design texts, the model improves the efficiency of evaluation and significantly improves the accuracy and depth of evaluation.

In terms of technology and market evaluation, through the fusion processing of multimodal data, the AI evaluation model can demonstrate an evaluation capability that is remarkably close to that of human experts or even exceeds that of humans in cases. For example, in the fourth part of the study, through a human-machine comparison experiment (Study 4), the study found that the consistency of the AI model with human experts in technology evaluation was as high as 73.58%. In market evaluation, it also reached 72.05%. These data thoroughly verify the effectiveness and reliability of the AI model in practical applications.

11.1.2 Potential for Improving Design Efficiency and Innovation of AI

The application of AI has significantly improved the efficiency of the design process. AI can quickly analyze and process design solutions in the data processing and preliminary evaluation stages, providing designers with immediate and valuable feedback. This capability not only shortens the design cycle but also improves the quality of the design. Through the second part of the study (Study 2), this thesis constructed a multimodal dataset containing more than thousands of design projects. The AI evaluation model can complete the preliminary evaluation of these design projects in a brief time, improving the efficiency of the evaluation work.

In addition, AI has also shown exciting potential in promoting design innovation. By analyzing historical design data and market trends, AI can reveal potential innovations that still need to be fully utilized in design, helping designers break through traditional design thinking and explore innovative design possibilities. In the third part of the study, Study 3 trained the AI model through deep learning algorithms, enabling it to identify and recommend innovative design concepts and methods, which designers have widely recognized and praised in practical applications.

11.1.3 Evaluation Model of Human-computer Collaboration

This study found that the human-machine collaborative evaluation model can effectively combine human designers' creativity and AI's data processing capabilities to provide a more comprehensive and objective design evaluation. Through experimental verification (Study 5), the model performs well in multiple evaluation dimensions, especially when dealing with complex and subjective evaluation tasks, and can provide results that match or even better

than human evaluation. For example, when dealing with the evaluation of design aesthetics and user experience, the AI model can extract valuable insights from a large amount of user feedback, helping designers understand user needs from a broader perspective and at a deeper level, as shown in Figure 11.1.

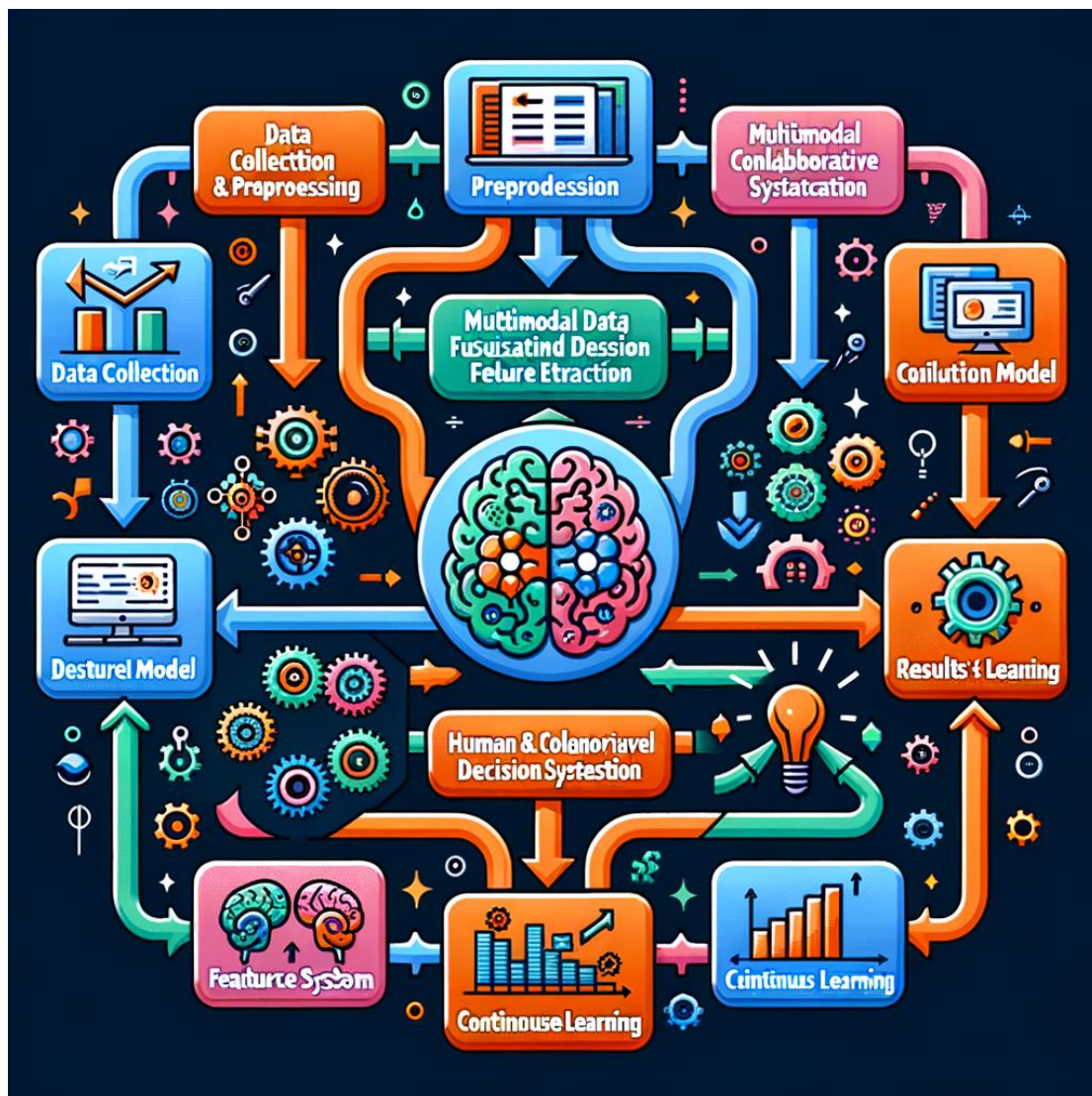


Figure 11.1 The Human-computer Collaborative Evaluation Framework

Figure 11.1 is generated by DALL-E-3. The prompts are generated by GPT-4-Turbo-128k based on part of the thesis content. The prompts are as follows: Flowchart illustrating an AI-assisted product design evaluation

framework. Start with 'Data Collection & Preprocessing' module, depicted with database and file icons, leading to 'Multimodal Data Fusion & Feature Extraction' module represented by overlapping circles and gears. Follow with 'AI Evaluation Model' module, showing a brain icon. Beside it, display 'Human-Machine Collaborative Decision Support System' with handshake icons. Next, link to 'Results Visualization & Feedback' module with chart icons, and end with 'Continuous Learning & Optimization' module, symbolized by an upward trend arrow. Use vibrant colors, clear labels, and directional arrows to indicate flow between modules.

Overall, this study demonstrates the practical application potential of AI in product design evaluation and proposes the broad application prospects of multimodal AI evaluation models in future design processes. Future research can further explore the application of AI technology in distinctive design fields (such as industrial design, fashion design, etc.) on this basis, optimize the performance of AI models, and improve their accuracy and applicability in design evaluation. At the same time, strengthening the research on the human-computer interaction interface in the AI evaluation process and improving the user-friendliness and transparency of AI systems are the keys to achieving the widespread application of AI technology in the design field. Through these efforts, AI technology will better serve designers, promote design innovation and efficiency, and achieve in-depth application and development in the design field.

11.2 Theoretical and Practical Contributions

This study has contributed both theoretical and practical, providing new insights and methodological frameworks for AI-assisted product design

evaluation. The following will discuss this study's theoretical contributions and practical significance in detail.

11.2.1 Theoretical Contributions

The main contributions of this study in design evaluation theory include exploring the application of multimodal deep learning technology in product concept design evaluation and providing an AI evaluation model that can simultaneously analyze product visual and text features. In addition, the study proposed a new human-computer interaction analysis framework, clarified the separate roles of human intelligence and artificial intelligence in design evaluation, and emphasized the importance of human-computer collaboration. Finally, the study focused on the transparency and interpretability of AI models, explored methods to improve the interpretability of AI evaluation results, and provided a theoretical basis for building responsible AI systems. These contributions have jointly promoted the in-depth application of AI technology in design evaluation and provided new methods and tools for innovation and product design optimization.

11.2.1.1 Application of Multimodal Deep Learning in Design Evaluation

This study is the first to systematically explore the application of multimodal deep learning in product design evaluation, providing innovative ideas for building a more comprehensive and accurate evaluation model. As described in Chapter 7, the ResNet18+BERT model we developed achieved excellent performance in multiple evaluation dimensions by fusing image and text information. This finding confirms the effectiveness of research results on multimodal machine translation in design (Zhou et al., 2018). It expands the

application of multimodal deep learning frameworks in activity and context recognition(Radu et al., 2018).

The research in this paper shows that multimodal deep learning models can more comprehensively capture all aspects of product design, including visual aesthetics, functional description, and market positioning. This method breaks through the limitations of traditional single-modality evaluation and provides a richer and more accurate information basis for design evaluation. This theoretical contribution promotes AI's application in the design field and provides new perspectives and methods for interdisciplinary research.

11.2.1.2 Research on AI Explainability in Design Evaluation

This study explores the interpretability of AI models in design evaluation and provides a theoretical basis for enhancing the credibility and acceptability of AI-assisted evaluation. As described in Chapter 9, designers' trust in AI evaluation results is highly correlated with their understanding of AI's work. This finding extends the application of the AI solution credibility framework proposed by Vianello et al. (2023) in the design field.

11.2.1.3 Establishment of Human-computer Collaboration Evaluation Framework

This study reveals the complementary advantages of AI models and human experts in design evaluation through empirical research, laying the foundation for establishing a more effective human-machine collaborative evaluation framework. Chapters 8 and 9 show that AI models perform well in objective, quantifiable evaluation dimensions (such as technology and

market), while human experts have advantages in subjective, creative evaluation. This finding echoes the research on designer confidence and AI collaboration effects (Chong et al., 2022) and expands the collaborative model of AI in the design process proposed by Figoli et al.(2022).

The human-computer collaborative evaluation framework proposed in this study emphasizes the respective advantages of AI and human experts. It explores how to integrate these advantages to effectively achieve the best evaluation results. This framework not only considers the accuracy and efficiency of the evaluation but also focuses on the interpretability and innovation of the evaluation process. The research results provide innovative ideas for solving the limitations of AI in engineering design proposed by Khanolkar et al. (2021).

This study proposes a series of methods to improve the interpretability of AI models, including developing visual explanation tools and implementing user education and training. These methods help designers better understand and accept AI evaluation results and pave the way for the broader application of AI in the creative field. This theoretical contribution provides a new perspective and method for solving the AI model interpretability challenge that Herskovits (2021)proposed.

11.2.2 Practical Significance

The AiDE system developed in this study provides designers with a practical multimodal AI-assisted evaluation tool that can simultaneously analyze a product's visual and textual features, improving the evaluation's accuracy and comprehensiveness. The results show that the application of AI technology

optimizes the product design innovation process, significantly improves design efficiency and innovation capabilities, and provides new methods and tools for product conceptualization, development, and optimization. In addition, this study provides an essential reference for design education, emphasizes the necessity of integrating AI technology into design courses, and helps to cultivate students' AI literacy and human-computer collaboration capabilities. Through these practical applications, this study promotes technological innovation in the design industry and provides essential guidance and inspiration for the widespread application of AI in the creative field.

The AiDE system developed in this study has demonstrated significant application advantages in the field of standardized product design. Its multimodal data fusion architecture (ResNet18+BERT) can effectively integrate product visual features and semantic descriptions, and achieve rapid quantitative verification of evaluation indicators for highly structured design scenarios such as functional products. The AiDE system has improved the efficiency of product design evaluation and provided a methodological framework for the generalization application of AI driven design evaluation tools in scenarios.

11.2.2.1 Providing Practical AI-assisted Evaluation Tools

The AI-assisted product design evaluation system (AiDE) developed in this study provides a practical evaluation tool for designers and enterprises. As described in Chapter 8, the AiDE system integrates advanced concepts such as multi-role user management, multimodal data processing, and human-computer collaborative evaluation, which can effectively improve the

efficiency and accuracy of design evaluation. The practical application value of this tool is consistent with the importance of AI in product development and responds to the empirical analysis of the adoption of AI-driven design tools (Chuyen, 2023; Manavis, 2023).

The development and application of the AiDE system provide a direct and available solution for design practice, which helps to improve design efficiency and product quality. By integrating AI technology with traditional design evaluation methods, the system provides new possibilities for innovation and development in the design field.

11.2.2.2 Optimizing the Product Design Innovation Process

The findings of this study provide important references for optimizing the product design innovation process. Through AI-assisted evaluation, designers can evaluate design concepts more quickly and accurately, accelerating the product development cycle. This echoes how AI changes the innovation and design process proposed by Verganti et al. (2020). This study further confirms the potential of AI in promoting design innovation and improving product market competitiveness.

This study shows that AI can help designers better understand and predict user needs, creating more innovative and market-competitive products. This data-driven design approach not only improves design efficiency but also enhances product technical and market adaptability, consistent with the findings of Zhang et al. (2022) on consumer responses to AI-designed products.

11.2.2.3 Promote the Application of AI Technology in the Design Industry

This study provides an important reference and impetus for the in-depth application of AI technology in the design industry. By systematically exploring the application of AI in product design evaluation, we verified the effectiveness of AI technology and revealed its potential to improve design quality and promote innovation. This echoes the review study of AI applications in engineering design by Yüksel et al. (2023) and expands the scope and depth of AI applications in the design field.

This study provides an empirical study of AI applications in the design industry, which helps promote the broader and deeper application of AI technology in the design industry. This not only helps to improve the innovation and competitiveness of the design industry but also provides new directions and content for design education.

11.2.2.4 Provide Reference for Design Education

As AI technology becomes more prevalent in design, future design education needs to incorporate AI knowledge and skills into the curriculum. This study provides important empirical references for design education. As described in Chapter 9, designers' familiarity with AI tools is closely related to their willingness to use and trust them. This finding highlights the importance of incorporating AI-related content into design education, which echoes the discussion on integrating emerging technologies into future designer education (McCardle, 2002).

This study provides specific guidance for design education, including how to cultivate students' AI literacy, integrate AI tools into design practice, and cultivate students' human-computer collaboration capabilities. These suggestions will help cultivate talents that meet the future design industry's needs and promote design education's innovation and development.

In summary, this study has made important contributions both in theory and practice. In theory, this study has promoted the application of multimodal deep learning in design evaluation, established a human-computer collaborative evaluation framework, and deepened the research on AI explainability in design evaluation. In practice, this study provides practical AI-assisted evaluation tools, provides guidance for optimizing the product design innovation process, promotes the in-depth application of AI technology in the design industry, and provides important references for design education. These contributions promote the development of AI-assisted product design evaluation and provide new perspectives for the future development of design practice and education.

11.3 Summary of Conclusion

As data, algorithms, and computing power continue to advance, AI will become an increasingly integral part of the creative and evaluation process. Expansions and improvements include the following aspects.

- Integration with emerging technologies

Future research could explore the integration of AI with innovative technologies such as AR to create more immersive and interactive design and evaluation experiences.

- Personalized and user-centered AI design

AI can help tailor designs to suit user preferences and needs by analyzing user feedback and behavioral patterns, potentially revolutionizing how products are conceptualized and marketed.

- Ethical and responsible use of AI

As AI plays a more significant role in design, addressing ethical issues such as data privacy, bias in AI algorithms, and the impact of AI on employment in the design sector will be crucial.

This thesis systematically explored the potential of AI in product design evaluation and developed and validated a multimodal AI-assisted evaluation system. The results show that AI technology has exciting potential to improve the efficiency and objectivity of design evaluation, especially when dealing with large-scale data and multi-dimensional evaluation. However, AI still faces challenges when dealing with highly subjective design innovation and aesthetic evaluation and needs to be combined with human professional judgment.

In the future, as AI technology continues to advance, we can expect the emergence of more intelligent and more transparent AI evaluation systems. These systems will better understand and evaluate complex design concepts,

providing designers with more valuable insights and suggestions. At the same time, the human-machine collaboration model will be further optimized, realizing the complementary advantages of AI and human designers, and promoting the dual improvement of design innovation and efficiency.

However, while embracing AI technology, we must also be vigilant about its potential ethical risks, such as data privacy, algorithmic bias, etc. Future research should focus on building responsible AI systems to ensure that technological progress and ethical considerations are equally important.

In short, integrating AI and design opens a new era of opportunities and challenges. Through continuous exploration and innovation, we are expected to achieve a change in the design evaluation model assisted by AI, bringing users a better and more innovative product experience. The development of this field will not only promote the progress of the design industry but also provide important reference and inspiration for the widespread application of AI in the creative field.

Finally, this thesis highlights AI's benefits in enhancing product design evaluation and lays the foundation for future advancements. As AI continues to advance, its integration into product design will enhance existing practices and unlock new creative potential, driving innovation and efficiency in unprecedented ways. This journey towards the future of AI integration has just begun, and there is great promise both in theoretical exploration and practical application.

References

- Abbasi, R., & Esmaili, M. (2024). Identifying the Applications of Artificial Intelligence in Online Marketing. In *Smart and Sustainable Interactive Marketing* (pp. 32-54). IGI Global.
<https://doi.org/10.4018/979-8-3693-1339-8.ch003IEE>
- Agrawal, J. (2010). Questionnaire Design. *Wiley International Encyclopedia of Marketing* (eds J. Sheth and N. Malhotra).
<https://doi.org/10.1002/9781444316568.wiem02013>
- Ahmad, M. F., Husin, N. A. A., Ahmad, A. N. A., Abdullah, H., Wei, C. S., & Nawi, M. N. M. (2022). Digital Transformation: An Exploring Barriers and Challenges Practice of Artificial Intelligence in Manufacturing Firms in Malaysia. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 29(1), 110–117. <https://doi.org/10.37934/araset.29.1.110117>
- Ahufinger, S., Balugo, P., González, M. M., Pequeño, E., González, H. L., & Herrero, P. (2019). A User-Centered Smartphone Application for Wireless EEG and Its Role in Epilepsy. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(6), 43. <https://doi.org/10.9781/ijimai.2019.06.004>
- Al-Sa'di, A., & Miller, D. G. (2023). Exploring the Impact of Artificial Intelligence Language Model ChatGPT on the User Experience. *International Journal of Technology Innovation and Management (Ijtim)*, 3(1), 1–8.
<https://doi.org/10.54489/ijtim.v3i1.195>
- Alufaisan, Y., Marusich, L. R., Bakdash, J. Z., Zhou, Y., & Kantarcioglu, M. (2021). Does explainable artificial intelligence improve human decision-making?. *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 35, No. 8, pp. 6618-6626). <https://doi.org/10.48550/arxiv.2006.11194>
- Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019, August). Guidelines for human-AI interaction. *Conference on Human Factors in Computing Systems - Proceedings*.
<https://doi.org/10.1145/3290605.3300233>
- Anggraini, S. F. (2022). The Validity of Interactive Multimedia Based on Cognitive Conflict on Elasticity Materials Using Adobe Animate CC 2019. *Journal of Social Sciences*, 8(1), 13–22. <https://doi.org/10.37134/ejoss.vol8.1.2.2022>
- Antoun, C., Katz, J., Argueta, J., & Wang, L. (2017). Design Heuristics for Effective Smartphone Questionnaires. *Social Science Computer Review*, 36(5), 557–574.
<https://doi.org/10.1177/0894439317727072>
- As, I., Pal, S., & Basu, P. (2018). Artificial Intelligence in Architecture: Generating Conceptual Design via Deep Learning. *International Journal of Architectural Computing*, 16(4), 306–327. <https://doi.org/10.1177/1478077118800982>

- Ayık, Z. (2023). Assessment of Creativity in Artifacts Designed by Gifted Students: A Social Semiotic Multimodal Perspective. *Bogazici University Journal of Education*, 40-2(1), 75-100. <https://doi.org/10.52597/buje.1109543>
- Bakalos, N., Voulodimos, A., Doulamis, N., Ostfeld, A., Salomons, E., Caubet, J., Jiménez, V., & Li, P. (2019). Protecting Water Infrastructure from Cyber and Physical Threats: Using Multimodal Data Fusion and Adaptive Deep Learning to Monitor Critical Systems. *IEEE Signal Processing Magazine*, 36(2), 36–48. <https://doi.org/10.1109/msp.2018.2885359>
- Bérubé, M., Giannelia, T., & Vial, G. (2021). Barriers to the Implementation of AI in Organizations: Findings from a Delphi Study. *Proceedings of the 54th Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/hicss.2021.805>
- Bogers, M., & Horst, W. (2013). Collaborative Prototyping: Cross-Fertilization of Knowledge in Prototype-Driven Problem Solving. *Journal of Product Innovation Management*, 31(4), 744–764. <https://doi.org/10.1111/jpim.12121>
- Boiani, J. A. M., Sandnes, F. E., Paschoarelli, L. C., & Medola, F. O. (2020). Satisfaction of aged users with mobility assistive devices: A preliminary study of conventional walkers. *Proceedings of the 2nd International Conference on Human Systems Engineering and Design (IHSED2019)*, Munich, Germany (pp. 742-746). Springer International Publishing. https://doi.org/10.1007/978-3-030-27928-8_113
- Botros, C. R., Mansour, Y., & Eleraky, A. (2023). Architecture Aesthetics Evaluation Methodologies of Humans and Artificial Intelligence. *Msa Engineering Journal*, 2(2), 450–462. <https://doi.org/10.21608/msaeng.2023.291897>
- Boursalie, O., Samavi, R., & Doyle, T. E. (2022). Evaluation methodology for deep learning imputation models. *Experimental Biology and Medicine*, 247(22), 1972–1987. <https://doi.org/10.1177/15353702221121602>
- Bulbul, M. F., Ullah, A., Ali, H., & Kim, D. (2022). A Deep Sequence Learning Framework for Action Recognition in Small-Scale Depth Video Dataset. *Sensors*, 22(18), 6841. <https://doi.org/10.3390/s22186841>
- Bulchand-Gidumal, J., Secin, E. W., O'Connor, P., & Buhalis, D. (2024). Artificial intelligence's impact on hospitality and tourism marketing: exploring key themes and addressing challenges. *Current Issues in Tourism*, 27(14), 2345–2362. <https://doi.org/10.1080/13683500.2023.2229480>
- Cai, J. (2024). Exploration of integrated teaching model for undergraduate and graduate students -- taking the course of questionnaire survey design and data analysis methods as an example. *SHS Web of Conferences*, 187, 01032. <https://doi.org/10.1051/shsconf/202418701032>
- Cai, J. (2024). Multimodal Graph Neural Architecture Search Under Distribution Shifts. *Proceedings of the Aaai Conference on Artificial Intelligence*, 38(8),

8227–8235. <https://doi.org/10.1609/aaai.v38i8.28663>

- Carvajal-Arango, R., Zuluaga-Holguín, D., & Mejía-Gutiérrez, R. (2014). A Systems-Engineering Approach for Virtual/Real Analysis and Validation of an Automated Greenhouse Irrigation System. *International Journal on Interactive Design and Manufacturing (Ijidem)*, 10(4), 355–367. <https://doi.org/10.1007/s12008-014-0243-2>
- Celinska, K. (2021). Evaluating a Functional Family Therapy Intervention: Challenges and Lessons Learned. *Research on Social Work Practice*, 31(4), 360–366. <https://doi.org/10.1177/1049731520985604>
- Ceschin, F., & Gaziulusoy, I. (2016). Evolution of design for sustainability: From product design to design for system innovations and transitions. *Design Studies*, 47, 118–163. <https://doi.org/10.1016/j.destud.2016.09.002>
- Chang, H.-Y. (2024). Exploring Metacognitive Processes in Design Ideation with Text-to-Image AI Tools. *Proceedings of the Design Society*, 4, 915–924. <https://doi.org/10.1017/pds.2024.94>
- Charmaz, K. (2015). Grounded theory. *Qualitative Psychology: A Practical Guide to Research Methods*, 3, 53–84.
- Chen, D., & Cheng, P. (2021). Development of Design System for Product Pattern Design Based on Kansei Engineering and BP Neural Network. *International Journal of Clothing Science and Technology*, 34(3), 335–346. <https://doi.org/10.1108/ijcst-04-2021-0044>
- Chen, J.-F. (2024). Designing the Future: A Case Study on Human-Ai Co-Innovation. *Creative Education*, 15(03), 474–494. <https://doi.org/10.4236/ce.2024.153028>
- Chen, R. (2016). Fuzzy Dual Experience-Based Design Evaluation Model for Integrating Engineering Design into Customer Responses. *International Journal on Interactive Design and Manufacturing (Ijidem)*, 10(4), 439–458. <https://doi.org/10.1007/s12008-016-0310-y>
- Chen, S. (2021). Embracing Multimodal Data in Multimedia Data Analysis. *IEEE Multimedia*, 28(3), 5–7. <https://doi.org/10.1109/mmul.2021.3104911>
- Choi, W. (2023). Developing an AI-based Automated Fashion Design System: Reflecting the Work Process of Fashion Designers. *Fashion and Textiles*, 10(1). <https://doi.org/10.1186/s40691-023-00360-w>
- Chong, L., Kotovsky, K., & Cagan, J. (2022). Are Confident Designers Good Teammates to Artificial Intelligence?: A Study of Self-Confidence, Competence, and Collaborative Performance. *Proceedings of the Design Society*, 2, 1531–1540. <https://doi.org/10.1017/pds.2022.155>
- Christou, P. A. (2023). The Use of Artificial Intelligence (AI) in Qualitative Research for Theory Development. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715/2023.6536>

- Chumiran, M. H., & Zainal Abidin, S. (2021). Design Pedagogy: Pictographic Design Artefacts Perceived Artificial Intelligence Elements for Product Designers. *Environment-Behaviour Proceedings Journal*, 6(SI4), 57–62. <https://doi.org/10.21834/ebpj.v6isi4.2901>
- Chuyen, N. T. H. (2023). An Empirical Analysis of Predictors of AI-Powered Design Tool Adoption. *TEM Journal*, 1482–1489. <https://doi.org/10.18421/tem123-28>
- Cornet, V. P., Toscos, T., Bolchini, D., Ghahari, R. R., Ahmed, R., Daley, C., Mirro, M. J., & Holden, R. J. (2020). Untold Stories in User-Centered Design of Mobile Health: Practical Challenges and Strategies Learned From the Design and Evaluation of an App for Older Adults With Heart Failure. *Jmir Mhealth and Uhealth*, 8(7), e17703. <https://doi.org/10.2196/17703>
- Creusen, M. E. H., & Schoormans, J. (2004). The Different Roles of Product Appearance in Consumer Choice. *Journal of Product Innovation Management*, 22(1), 63–81. <https://doi.org/10.1111/j.0737-6782.2005.00103.x>
- Creusen, M. E. H., Veryzer, R. W., & Schoormans, J. (2010). Product Value Importance and Consumer Preference for Visual Complexity and Symmetry. *European Journal of Marketing*, 44(9/10), 1437–1452. <https://doi.org/10.1108/03090561011062916>
- Davis, N., Hsiao, C. P., Yashraj Singh, K., Li, L., & Magerko, B. (2016). Empirically studying participatory sense-making in abstract drawing with a co-creative cognitive agent. *Proceedings of the 21st International Conference on Intelligent User Interfaces* (pp. 196-207). <https://doi.org/10.1145/2856767.2856795>
- De Azevedo, A. P., Tives, H. A., & Canedo, E. D. (2022). Guide for artificial intelligence ethical requirements elicitation—re4ai ethical guide. *Proceedings of the 55th Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/hicss.2022.677>
- Denton, H., & McDonagh, D. (2003). Using focus group methods to improve students' design project research in schools: Drawing parallels from action research at undergraduate level. *International Journal of Technology and Design Education*, 13, 129-144. .
- Devlin, J. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. <https://doi.org/10.48550/arxiv.1810.04805>
- DeVaughn, P., & Stefaniak, J. E. (2020). An Exploration of the Challenges Instructional Designers Encounter While Conducting Evaluations. *Performance Improvement Quarterly*, 33(4), 443–470. <https://doi.org/10.1002/piq.21332>
- Dhillon, P. S., Molaei, S., Li, J., Golub, M., Zheng, S., & Robert, L. P. (2024). Shaping Human-AI Collaboration: Varied Scaffolding Levels in Co-writing with Language Models. *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-18). <https://doi.org/10.1145/3613904.3642134>

- Dlugatch, R., Georgieva, A., & Kerasidou, A. (2023). Trustworthy artificial intelligence and ethical design: public perceptions of trustworthiness of an AI-based decision-support tool in the context of intrapartum care. *BMC Medical Ethics*, 24(1). <https://doi.org/10.1186/s12910-023-00917-w>
- Donald, N. (2013). *The design of everyday things: Revised and expanded edition*. 56(3), 785. <https://doi.org/10.1353/tech.2015.0104>
- Dosovitskiy, A. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*. <http://arxiv.org/abs/2010.11929>
- ERTÜRK, M., & ÜZÜMCÜ, N. E. (2022). Third Eye in Evaluation Processes of Visual Communication Design Education: An Artificial Intelligence Model. *Sanat Ve Tasarım Dergisi*, 12(1), 191–202. <https://doi.org/10.20488/sanattasarim.1133850>
- Fan, X., Li, X., Song, L., Zhang, Y., Liu, M., & Miao, Q. (2022). Usability Evaluation of AR Human-machine Interface Based on User Experience. *Proceedings - 2022 International Conference on Virtual Reality, Human-Computer Interaction and Artificial Intelligence, VRHCIAI 2022*, 23–29. <https://doi.org/10.1109/VRHCIAI57205.2022.00011>
- Fang, J. (2023). Research on the Design of Business Models and Transformation Management of New Entrepreneurial Ventures Driven by Artificial Intelligence. *BCP Business & Management*, 49, 36–41. <https://doi.org/10.54691/bcpbm.v49i.5383>
- Feuston, J. L., & Brubaker, J. R. (2021). Putting Tools in Their Place: The Role of Time and Perspective in Human-AI Collaboration for Qualitative Analysis. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2). <https://doi.org/10.1145/3479856>
- Figoli, F. A., Rampino, L., & Mattioli, F. (2022). AI in the design process: training the human-AI collaboration. *Proceedings of the 24th International Conference on Engineering and Product Design Education (E&PDE 2022)* (pp. 1-6). <https://doi.org/10.35199/epde.2022.61>
- Fogliato, R., Chappidi, S., Lungren, M., Fisher, P., Wilson, D., Fitzke, M., ... & Nushi, B. (2022). Who goes first? Influences of human-AI workflow on decision making in clinical imaging. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1362-1374). <https://doi.org/10.48550/arxiv.2205.09696>
- Freelon, D. (2013). Recal oir: ordinal, interval, and ratio intercoder reliability as a web service. *International Journal of Internet Science*, 8(1), 10-16.
- Fujii, G., Hamada, K., Ishikawa, F., Masuda, S., Matsuya, M., Myojin, T., Nishi, Y., Ogawa, H., Toku, T., Tokumoto, S., Tsuchiya, K., & Ujita, Y. (2020). Guidelines for Quality Assurance of Machine Learning-Based Artificial Intelligence.

- International Journal of Software Engineering and Knowledge Engineering*, 30(11–12), 1589–1606. <https://doi.org/10.1142/S0218194020400227>
- Gao, J., Choo, K. T. W., Cao, J., Lee, R. K. W., & Perrault, S. (2023). CoAICoder: Examining the effectiveness of AI-assisted human-to-human collaboration in qualitative analysis. *ACM Transactions on Computer-Human Interaction*, 31(1), 1–38. <http://arxiv.org/abs/2304.05560>
- Gao, W., Pei, Y., Liang, H., Lv, J., Chen, J., & Zhong, W. (2021). Multimodal AI System for the Rapid Diagnosis and Surgical Prediction of Necrotizing Enterocolitis. *IEEE Access*, 9, 51050–51064. <https://doi.org/10.1109/access.2021.3069191>
- Gemrot, J., Černý, M., & Brom, C. (2014, March). Why you should Empirically Evaluate your AI Tool-From SPOSH to yaPOSH. *International Conference on Agents and Artificial Intelligence* (Vol. 2, pp. 461–468). SCITEPRESS. <https://doi.org/10.5220/0004818604610468>
- Gezici, B., & Tarhan, A. K. (2022). Systematic literature review on software quality for AI-based software. *Empirical Software Engineering*, 27(3). <https://doi.org/10.1007/s10664-021-10105-2>
- Glaser, B. G., Ekstrom, H., Esseveld, J., Hovelius, B., Raffanti, M. A., Fernandez, W. D., & Lehmann, H. (2005). The Grounded Theory Review: An international journal Sociology Press Staying Open. *The Grounded Theory Review* (Vol. 5, Issue 1).
- Gmeiner, F., Yang, H., Yao, L., Holstein, K., & Martelaro, N. (2023). *Exploring Challenges and Opportunities to Support Designers in Learning to Co-Create With AI-based Manufacturing Design Tools*. <https://doi.org/10.1145/3544548.3580999>
- Gonzalez-Hernandez, G., Sarker, A. A., O'Connor, K., & Savova, G. (2017). Capturing the Patient's Perspective: A Review of Advances in Natural Language Processing of Health-Related Text. *Yearbook of Medical Informatics*, 26(01), 214–227. <https://doi.org/10.1055/s-0037-1606506>
- Gordon, T. J., Wolfson, R. J., & Sahr, R. C. (1969). Methodology. *Futures.*, 1(6), 552. [https://doi.org/10.1016/S0016-3287\(69\)80045-5](https://doi.org/10.1016/S0016-3287(69)80045-5)
- Haase, J., Wiedmann, K., Bettels, J., & Labenz, F. (2018). How to Best Promote My Product? Comparing the Effectiveness of Sensory, Functional and Symbolic Advertising Content in Food Marketing. *British Food Journal*, 120(8), 1792–1806. <https://doi.org/10.1108/bfj-01-2018-0058>
- Hao, Z., Li, Z., Dang, X., Ma, Z., & Liu, G. (2022). MM-LMF: A Low-Rank Multimodal Fusion Dangerous Driving Behavior Recognition Method Based on FMCW Signals. *Electronics*, 11(22), 3800. <https://doi.org/10.3390/electronics11223800>
- Hassan, M. F., Saman, M. Z. M., Sharif, S., & Omar, B. (2012). Integration of Morphological Analysis Theory and Artificial Neural Network Approach for

- Sustainable Product Design: A Case Study of Portable Vacuum Cleaner. *International Journal of Sustainable Manufacturing*, 2(4), 293.
<https://doi.org/10.1504/ijsm.2012.048585>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778). <https://doi.org/10.1109/cvpr.2016.90>
- Hegazy, M., & Saleh, A. I. (2023). Evolution of AI Role in Architectural Design: Between Parametric Exploration and Machine Hallucination. *Msa Engineering Journal*, 2(2), 262–288. <https://doi.org/10.21608/msaeng.2023.291873>
- Hendriana, B., Apoko, T. W., Hadi, W., Handayani, I., Supandi, S., & E, N. (2023). The Independent Learning-Independent Campus Policy: Students' Awareness of Mathematics Education. *Al-Ishlah Jurnal Pendidikan*, 15(2), 1238–1245.
<https://doi.org/10.35445/alishlah.v15i2.2302>
- Herskovits, E. H. (2021). Artificial Intelligence in Molecular Imaging. *Annals of Translational Medicine*, 9(9), 824–824. <https://doi.org/10.21037/atm-20-6191>
- Holmes, G., Dawson, O., Waltho, D., Beaty, J., & Newnes, C. (1998). User Views of Two Psychiatric Day Hospitals. *Psychiatric Bulletin*, 22(6), 362–364.
<https://doi.org/10.1192/pb.22.6.362>
- Hou, Y., & Xu, X. (2021). Research on Art Design and Application of Indoor Environment Based on Artificial Intelligence. *E3s Web of Conferences*, 275, 3036.
<https://doi.org/10.1051/e3sconf/202127503036>
- Hu, J. (2022). Product Design Method Based on Data Fusion and Transmission Based on Multimode Sensor. *Mobile Information Systems*, 2022, 1–10.
<https://doi.org/10.1155/2022/5709786>
- Hu, J. (2024). *Research on Innovative Design Method and Evaluation of Jewelry Based on CiteSpace&Playground AI*. <https://doi.org/10.3233/faia231424>
- Huang, S. (2022). Product Innovation Design Method Based on BP Neural Network. *Advances in Multimedia*, 2022, 1–6. <https://doi.org/10.1155/2022/6830892>
- Jha, A. K., Bradshaw, T., Buvat, I., Hatt, M., Kc, P., Liu, C., Obuchowski, N., Saboury, B., Slomka, P. J., Sunderland, J., Wahl, R. L., Yu, Z., Zuehlsdorff, S., Rahmim, A., & Boellaard, R. (2022). Nuclear Medicine and Artificial Intelligence: Best Practices for Evaluation (The RELAINCE Guidelines). *Journal of Nuclear Medicine*, 63(9), 1288–1299. <https://doi.org/10.2967/jnumed.121.263239>
- Jian, H., Shen, G., & Ren, X. (2021). Connotation Analysis and Paradigm Shift of Teaching Design Under Artificial Intelligence Technology. *International Journal of Emerging Technologies in Learning (Ijet)*, 16(05), 73.
<https://doi.org/10.3991/ijet.v16i05.20287>
- Jiang, N., Liu, X., Liu, H., Lim, E. T. K., Tan, C., & Gu, J. (2022). Beyond AI-powered Context-Aware services: The Role Of human–AI Collaboration. *Industrial Management & Data Systems*, 123(11), 2771–2802.

<https://doi.org/10.1108/imds-03-2022-0152>

- Jin, X., Wu, L., Song, C., Zhao, G., Chen, S., Chi, J., Peng, S., & Ge, S. (2018). Predicting Aesthetic Score Distribution Through Cumulative Jensen-Shannon Divergence. *Proceedings of the Aaai Conference on Artificial Intelligence*, 32(1). <https://doi.org/10.1609/aaai.v32i1.11286>
- Jones, R. A., Lubans, D. R., Morgan, P. J., Okely, A. D., Parletta, N., Wolfenden, L., Silva-Sanigorski, A. d., Gibbs, L., & Waters, E. (2014). School-Based Obesity Prevention Interventions: Practicalities and Considerations. *Obesity Research & Clinical Practice*, 8(5), e497–e510. <https://doi.org/10.1016/j.orcp.2013.10.004>
- Jylkäs, T., Augsten, A., & Miettinen, S. (2019). From Hype to Practic. *Conference Proceedings of the Academy for Design Innovation Management*, 2(1). <https://doi.org/10.33114/adim.2019.04.349>
- Kanchymalay, K., Krishnan, R., Arif, F., Amiruddin, S., Salam, S., & Hashim, U. R. (2013). The Extent of ERP Customization Towards User Satisfaction in Daily Operation for Manufacturing Companies. *Journal of Computers*, 8(7). <https://doi.org/10.4304/jcp.8.7.1788-1792>
- Kane, C., Rubio, D., & Trochim, W. (2013). Evaluating translational research. *Translational Medicine-What, Why and How: An International Perspective* (Vol. 3, pp. 110-119). Karger Publishers. <https://doi.org/10.1159/000343025>
- Kang, C., Valenzise, G., & Dufaux, F. (2020). Eva: An explainable visual aesthetics dataset. *Joint workshop on aesthetic and technical quality assessment of multimedia and media analytics for societal trends* (pp. 5-13). <https://doi.org/10.1145/3423268.3423590>
- Karimi, P., Grace, K., Maher, M. L., & Davis, N. (2018). Evaluating creativity in computational co-creative systems. *arXiv preprint arXiv:1807.09886*. <https://doi.org/10.48550/arxiv.1807.09886>
- Kato, T., & Tsuda, K. (2018). Quantitative Evaluation of Sensitivity in Confidential Car Exterior Design. *Artificial Intelligence Research*, 7(1), 34. <https://doi.org/10.5430/air.v7n1p34>
- Kersloot, M. G., van Putten, F. J. P., Abu-Hanna, A., Cornet, R., & Arts, D. L. (2020). Natural Language Processing Algorithms for Mapping Clinical Text Fragments onto Ontology Concepts: A Systematic Review and Recommendations for Future Studies. *Journal of Biomedical Semantics*, 11(1). <https://doi.org/10.1186/s13326-020-00231-z>
- Khanolkar, P. M., Gad, M. M., Liao, J. C., Hurst, A., & Olechowski, A. (2021). A Pilot Study on the Prevalence of Artificial Intelligence in Canadian Engineering Design Curricula. *Proceedings of the Canadian Engineering Education Association (Ceea)*. <https://doi.org/10.24908/pceea.vi0.14919>
- Kim, D. W., Jang, H. Y., Kim, K. W., Shin, Y., & Park, S. H. (2019). Design Characteristics of Studies Reporting the Performance of Artificial Intelligence

- Algorithms for Diagnostic Analysis of Medical Images: Results from Recently Published Papers. *Korean Journal of Radiology*, 20(3), 405.
<https://doi.org/10.3348/kjr.2019.0025>
- Kim, K. G. (2016). Book Review: Deep Learning. *Healthcare Informatics Research*, 22(4), 351. <https://doi.org/10.4258/hir.2016.22.4.351>
- Kitchenham, B. A., & Pfleeger, S. L. (2008). Personal opinion surveys. *Guide to advanced empirical software engineering* (pp. 63-92). London: Springer London.
https://doi.org/10.1007/978-1-84800-044-5_3
- Kong, F. (2020). Application of Artificial Intelligence in Modern Art Teaching. *International Journal of Emerging Technologies in Learning*, 15(13), 238.
<https://doi.org/10.3991/ijet.v15i13.15351>
- Kong, H. J., & Kim, Y. L. (2024). Application of artificial intelligence in dental crown prosthesis: a scoping review. *BMC Oral Health*, 24(1), 937.
<https://doi.org/10.21203/rs.3.rs-4493582/v1>
- Kreimeyer, K., Foster, M., Pandey, A., Arya, N., Halford, G., Jones, S. F., Forshee, R. A., Walderhaug, M., & Botsis, T. (2017). Natural Language Processing Systems for Capturing and Standardizing Unstructured Clinical Information: A Systematic Review. *Journal of Biomedical Informatics*, 73, 14–29.
<https://doi.org/10.1016/j.jbi.2017.07.012>
- Kuang, E., Jahangirzadeh Soure, E., Fan, M., Zhao, J., & Shinohara, K. (2023, April). Collaboration with conversational AI assistants for UX evaluation: Questions and how to ask them (voice vs. text). *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1-15). 1–15.
<https://doi.org/10.1145/3544548.3581247>
- Kuo, T. S., Halfaker, A. L., Cheng, Z., Kim, J., Wu, M. H., Wu, T., ... & Zhu, H. (2024, May). Wikibench: Community-driven data curation for ai evaluation on wikipedia. *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-24). <https://doi.org/10.1145/3613904.3642278>
- Kuzovkin, I., Vicente, R., Petton, M., Lachaux, J. P., Baciú, M., Kahane, P., ... & Aru, J. (2018). Activations of deep convolutional neural networks are aligned with gamma band activity of human visual cortex. *Communications biology*, 1(1), 107.
<https://doi.org/10.1101/133694>
- Lakhani, P., & Sundaram, B. (2017). Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology*, 284(2), 574–582.
<https://doi.org/10.1148/radiol.2017162326>
- Lange, E. M. (2024). Flow in the Age of AI: The Impact of AI-enabled Work Tools on IT Professionals and Design Recommendations (Dissertation). Retrieved from <https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-532349>
- Lankes, M., & Stockl, A. (2023). AI-Powered Game Design: Experts Employing

- ChatGPT in the Game Design Process. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, 24, 1-9.
<https://doi.org/10.55549/epstem.1406194>
- Latour, B. (2009). Spheres and networks. Two ways to reinterpret globalization. *Harvard Design Magazine*, (30), 138-144.
<https://sciencespo.hal.science/hal-01022658>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lee, S., Isaksen, A., Holmgård, C., & Togelius, J. (2021). Predicting Resource Locations in Game Maps Using Deep Convolutional Neural Networks. *Proceedings of the Aaai Conference on Artificial Intelligence and Interactive Digital Entertainment*, 12(2), 46–52. <https://doi.org/10.1609/aiide.v12i2.12893>
- Lepori, M. A., & Firestone, C. (2022). Can you hear me now? Sensitive comparisons of human and machine perception. *Cognitive Science*, 46(10), e13191.
<https://doi.org/10.48550/arxiv.2003.12362>
- Levman, J., Ewenson, B., Apaloo, J., Berger, D., & Tyrrell, P. N. (2023). Error Consistency for Machine Learning Evaluation and Validation with Application to Biomedical Diagnostics. *Diagnostics*, 13(7), 1315.
<https://doi.org/10.3390/diagnostics13071315>
- Li, M., Zhang, H., Xu, C., Yan, C., Liu, H., & Li, X. (2022). MFVC: Urban Traffic Scene Video Caption Based on Multimodal Fusion. *Electronics*, 11(19), 2999.
<https://doi.org/10.3390/electronics11192999>
- Li, L., & Zhang, W. (2019). Precision marketing driven by the internet supply chain in the new retail era. In *Fourth International Conference on Economic and Business Management (FEBM 2019)* (pp. 212-214). Atlantis Press.
<https://doi.org/10.2991/feb-19.2019.52>
- Linsey, J. S., Laux, J. P., Clauss, E., Wood, K. L., & Markman, A. B. (2007). Increasing innovation: A trilogy of experiments towards a design-by-analogy method. *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (Vol. 48043, pp. 145-159).
<https://doi.org/10.1115/detc2007-34948>
- Liu, C. (2022). Artificial Intelligence Interactive Design System Based on Digital Multimedia Technology. *Advances in Multimedia*, 2022, 1–12.
<https://doi.org/10.1155/2022/4679066>
- Li, F. (2015). Network collaborative modeling design of intelligent products in industrial design system CNC machine tools. *2015 International Conference on Automation, Mechanical Control and Computational Engineering* (pp. 1074-1079). Atlantis Press. <https://doi.org/10.2991/amcce-15.2015.148>
- Liu, K. long, Li, W., Yang, C. yuan, & Yang, G. (2019). Intelligent design of multimedia content in Alibaba. *Frontiers of Information Technology and Electronic*

- Engineering*, 20(12), 1657–1664. <https://doi.org/10.1631/FITEE.1900580>
- Liu, Y., & Kim, K. S. (2023). An artificial-intelligence-driven product design framework with a synergistic combination of Genetic Algorithm and Particle Swarm Optimization. *Soft Computing*, 27(23), 17621–17638. <https://doi.org/10.1007/s00500-023-09223-4>
- Lobo, A., Duarte, P., Carvalho, A. A. de S., Rodrigues, V. M. C. P., Monteiro, M. J., & Alves, H. (2013). The Association of Equity, Accessibility, and Price with Primary Healthcare User's Satisfaction. *Western Journal of Nursing Research*, 36(2), 191–208. <https://doi.org/10.1177/0193945913497830>
- Louie, J., Rhoads, C., & Mark, J. (2016). Challenges to Using the Regression Discontinuity Design in Educational Evaluations. *American Journal of Evaluation*, 37(3), 381–407. <https://doi.org/10.1177/1098214015621787>
- Lopez, A., & Garza, R. (2023). Consumer bias against evaluations received by artificial intelligence: the mediation effect of lack of transparency anxiety. *Journal of Research in Interactive Marketing*, 17(6), 831–847. <https://doi.org/10.1108/JRIM-07-2021-0192>
- Lu, F., Xu, Y., Xu, X., Jones, B., & Malamed, L. (2023). Exploring the Impact of User and System Factors on Human-AI Interactions in Head-Worn Displays. *Proceedings - 2023 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2023*, 109–118. <https://doi.org/10.1109/ISMAR59233.2023.00025>
- Lu, M.-X., Du, G.-Z., & Li, Z. (2022). Multimode Gesture Recognition Algorithm Based on Convolutional Long Short-Term Memory Network. *Computational Intelligence and Neuroscience*, 2022, 1–10. <https://doi.org/10.1155/2022/4068414>
- Ma, B., Hauer, R. J., & Xu, C. (2020). Effects of Design Proportion and Distribution of Color in Urban and Suburban Green Space Planning to Visual Aesthetics Quality. *Forests*, 11(3), 278. <https://doi.org/10.3390/f11030278>
- Ma, Y., Jin, J., Yu, W., Zhang, W., Xu, Z., & Ma, Q. (2018). How Is the Neural Response to the Design of Experience Goods Related to Personalized Preference? An Implicit View. *Frontiers in Neuroscience*, 12. <https://doi.org/10.3389/fnins.2018.00760>
- Madaio, M., Egede, L., Subramonyam, H., Vaughan, J., & Wallach, H. (2022). Assessing the Fairness of AI Systems: AI Practitioners' Processes, Challenges, and Needs for Support. *Proceedings of the Acm on Human-Computer Interaction*, 6(CSCW1), 1–26. <https://doi.org/10.1145/3512899>
- Madni, H. A. (2024). Exploiting Data Diversity in Multi-Domain Federated Learning. *Machine Learning Science and Technology*, 5(2), 25041. <https://doi.org/10.1088/2632-2153/ad4768>
- MAKSOU, N. A. E. (2024). Artificial Intelligence Applications in Green Architecture. *Fayoum University Journal of Engineering*, 7(2), 317–337. <https://doi.org/10.21608/fuje.2024.345049>

- Malu, G., Bapi, R. S., & Indurkha, B. (2017). Learning photography aesthetics with deep cnns. *arXiv preprint arXiv:1707.03981*.
<https://doi.org/10.48550/arxiv.1707.03981>
- Manavis, A. (2023). Artificial Intelligence in Product Development. *Forum A+p*, 27, 16–21. <https://doi.org/10.37199/f40002702>
- McCardle, J. R. (2002). The challenge of integrating ai & smart technology in design education. *International Journal of Technology and Design Education*, 12, 59-76.
- Miranda, S. (2024). User Satisfaction Regarding Regional Health Polyclinics Services in Bahia, Brazil. *Concilium*, 24(12), 338–352.
<https://doi.org/10.53660/clm-3635-24m32>
- M.Matter, N. (2024). Artificial Intelligence in Architecture: Integration into Architectural Design Process. *Engineering Research Journal*, 181(0), 1–16.
<https://doi.org/10.21608/erj.2024.344313>
- Moghaddam, F. B. (2024). Natural Language Processing in Knowledge-Based Support for Operator Assistance. *Applied Sciences*, 14(7), 2766.
<https://doi.org/10.3390/app14072766>
- Mondal, S., & Goswami, S. S. (2024). Rise of Intelligent Machines: Influence of Artificial Intelligence on Mechanical Engineering Innovation. *Spectrum of Engineering and Management Sciences*, 2(1), 46-55.
<https://doi.org/10.31181/sems1120244h>
- Moore, K. R. (2023). Respecting Tribal Voices in the Development of a Gestational Diabetes Risk Reduction Preconception Counseling Program for American Indian/Alaska Native Adolescent Females: A Qualitative Study. *BMC Pregnancy and Childbirth*, 23(1). <https://doi.org/10.1186/s12884-023-05850-9>
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2019). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics*, 26(4), 2141–2168. <https://doi.org/10.1007/s11948-019-00165-5>
- Mundher, R., Bakar, S. A., Al-Helli, M., Gao, H., Al-Sharaa, A., Yusof, M. J. M., Maulan, S., & Aziz, A. (2022). Visual Aesthetic Quality Assessment of Urban Forests: A Conceptual Framework. *Urban Science*, 6(4), 79.
<https://doi.org/10.3390/urbansci6040079>
- Nabilah, N. (2024). Redefining Assessment: Creating a Groundbreaking Prototype for Domain Affective in Project-Based Learning. *Contemporary Mathematics and Science Education*, 5(1), ep24005. <https://doi.org/10.30935/conmaths/14457>
- Neto, J. C., Gaieski, T., Amaral, A. M., & Colanzi, T. E. (2018). Quanti-qualitative analysis of a memetic algorithm to optimize product line architecture design. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, 2018-November*, 498–505. <https://doi.org/10.1109/ICTAI.2018.00083>
- Niraula, D., Cuneo, K. C., Dinov, I. D., Gonzalez, B. D., Jamaluddin, J. B., Jin, J. J., ...

- & El Naqa, I. (2024). Intricacies of Human-AI Interaction in Dynamic Decision-Making for Precision Oncology: A Case Study in Response-Adaptive Radiotherapy. *medRxiv*. <https://doi.org/10.1101/2024.04.27.24306434>
- Oh, S., Jung, Y., Kim, S., Lee, I., & Kang, N. (2019). Deep Generative Design: Integration of Topology Optimization and Generative Models. *Journal of Mechanical Design*, 141(11). <https://doi.org/10.1115/1.4044229>
- Ohueri, C. C., Enegbuma, W. I., & Habil, H. (2019). MyCREST Embedded Framework for Enhancing the Adoption of Green Office Building Development in Sarawak. *Built Environment Project and Asset Management*, 10(2), 215–230. <https://doi.org/10.1108/bepam-10-2018-0127>
- Okafor, E. J., & Akcay, E. (2024). Applying User-Centered Design Methods to Improve The Experience of the NHS APP. *Qeios*. <https://doi.org/10.32388/byaenm>
- Ørevik, S. J. H. (2023). Assessing Students' Multimodal Texts in the Subject of English: Synthesising Peers' and Teachers' Recognition of Semiotic Work. *Designs for Learning*, 15(1). <https://doi.org/10.16993/dfl.216>
- Pal, N. S. S., Raymahapatra, P., Paul, S., Dolui, S., Chaudhuri, A. K., & Das, S. (2023). A Novel Brain Tumor Classification Model Using Machine Learning Techniques. *International Journal of Engineering Technology and Management Sciences*, 7(2), 87–98. <https://doi.org/10.46647/ijetms.2023.v07i02.011>
- Park, S. H., & Han, K. (2018). Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. *Radiology*, 286(3), 800–809. <https://doi.org/10.1148/radiol.2017171920>
- Partlan, N., Kleinman, E., Howe, J., Ahmad, S., Marsella, S., & Seif El-Nasr, M. (2021, August). Design-driven requirements for computationally co-creative game AI design tools. In *Proceedings of the 16th International Conference on the Foundations of Digital Games* (pp. 1-12). <https://doi.org/10.1145/3472538.3472573>
- Peuter, S. De, Oulasvirta, A., & Kaski, S. (2023). Toward AI assistants that let designers design. *AI Magazine*, 44(1), 85–96. <https://doi.org/10.1002/aaai.12077>
- Piriyakul, I., Piriyakul, R., & Piriyakul, M. (2022). Analysis of Thai text from social media for mass customization. *Journal of Social Sciences*, 43(3), 611–618. <https://doi.org/10.34044/j.kjss.2022.43.3.11>
- Potnis, K. C., Ross, J. S., Aneja, S., Gross, C. P., & Richman, I. B. (2022). Artificial Intelligence in Breast Cancer Screening. *Jama Internal Medicine*, 182(12), 1306. <https://doi.org/10.1001/jamainternmed.2022.4969>
- Quan, H., Li, S., Zeng, C., Wei, H., & Hu, J. (2023). Big Data and AI-Driven Product Design: A Survey. *Applied Sciences (Switzerland)* (Vol. 13, Issue 16). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/app13169433>

- Quan, S. J., Park, J., Economou, A., & Lee, S. (2019). Artificial Intelligence-assisted Design: Smart Design for Sustainable City Development. *Environment and Planning B Urban Analytics and City Science*, 46(8), 1581–1599. <https://doi.org/10.1177/2399808319867946>
- Radu, V., Tong, C., Bhattacharya, S., Lane, N. D., Mascolo, C., Marina, M. K., & Kawsar, F. (2018). Multimodal Deep Learning for Activity and Context Recognition. *Proceedings of the Acm on Interactive Mobile Wearable and Ubiquitous Technologies*, 1(4), 1–27. <https://doi.org/10.1145/3161174>
- Rampersad, G. (2020). Robot will take your job: Innovation for an era of artificial intelligence. *Journal of Business Research*, 116, 68–74. <https://doi.org/10.1016/j.jbusres.2020.05.019>
- Ran, X., Shi, J., Chen, Y., & Jiang, K. (2022). Multimodal Neuroimage Data Fusion Based on Multikernel Learning in Personalized Medicine. *Frontiers in Pharmacology*, 13. <https://doi.org/10.3389/fphar.2022.947657>
- Rath, R. C., Baral, S. K., Singh, T., & Goel, R. (2022). Role of Artificial Intelligence and Machine Learning in Product Design and Manufacturing. 2022 *International Mobile and Embedded Technology Conference, MECON 2022*, 571–575. <https://doi.org/10.1109/MECON53876.2022.9752455>
- Rattray, J., & Jones, M. (2007). Essential Elements of Questionnaire Design and Development. *Journal of Clinical Nursing*, 16(2), 234–243. <https://doi.org/10.1111/j.1365-2702.2006.01573.x>
- Rezwana, J., & Maher, M. L. (2022). Understanding user perceptions, collaborative experience and user engagement in different human-AI interaction designs for co-creative systems. *Proceedings of the 14th Conference on Creativity and Cognition* (pp. 38-48). <https://doi.org/10.48550/arxiv.2204.13217>
- Rezwana, J., & Maher, M. L. (2022). Identifying ethical issues in ai partners in human-ai co-creation. *arXiv preprint arXiv:2204.07644*. <https://doi.org/10.48550/arxiv.2204.07644>
- Rosen, R. K., Gainey, M., Nasrin, S., Garbern, S. C., Lantini, R., Elshabassi, N., Sultana, S., Hasnin, T., Alam, N. H., Nelson, E. J., & Levine, A. C. (2023). Use of Framework Matrix and Thematic Coding Methods in Qualitative Analysis for mHealth: The FluidCalc app. *International Journal of Qualitative Methods*, 22. <https://doi.org/10.1177/16094069231184123>
- Ruiz-Rojas, L. I. (2023). Empowering Education with Generative Artificial Intelligence Tools: Approach With an Instructional Design Matrix. *Sustainability*, 15(15), 11524. <https://doi.org/10.3390/su151511524>
- Saadi, J., & Yang, M. C. (2023). Observations on the Implications of Generative Design Tools on Design Process and Designer Behaviour. *Proceedings of the Design Society*, 3, 2805–2814. <https://doi.org/10.1017/pds.2023.281>
- Said, A. B., Mohamed, A., Elfouly, T., Harras, K., & Wang, Z. J. (2017). Multimodal

- deep learning approach for joint EEG-EMG data compression and classification. *2017 IEEE wireless communications and networking conference (WCNC)* (pp. 1-6). IEEE. <https://doi.org/10.1109/wcnc.2017.7925709>
- Saidani, M., Kim, H., & Yannou, B. (2021). Can machine learning tools support the identification of sustainable design leads from product reviews? Opportunities and challenges. *International design engineering technical conferences and computers and information in engineering conference* (Vol. 85383, p. V03AT03A005). American Society of Mechanical Engineers.
- Sanabria-Z, J., & Olivo, P. G. (2024). AI platform model on 4IR megatrend challenges: complex thinking by active and transformational learning. *Interactive Technology and Smart Education*. <https://doi.org/10.1108/ITSE-07-2023-0145>
- Satwiko, P., & Michelle, B. (2022). From Science Fiction to Science Facts. *International Webinar on Digital Architecture 2021 (IWEDA 2021)*, 278-284. <https://doi.org/10.2991/assehr.k.220703.050>
- Schoonderwoerd, T., Zoelen, E. M. v., Bosch, K. V. d., & Neerincx, M. A. (2022). Design Patterns for Human-Ai Co-Learning: A Wizard-of-Oz Evaluation in an Urban-Search-and-Rescue Task. *International Journal of Human-Computer Studies*, 164, 102831. <https://doi.org/10.1016/j.ijhcs.2022.102831>
- Selva, D., Cameron, B., & Crawley, E. F. (2014). A Rule-Based Method for Scalable and Traceable Evaluation of System Architectures. *Research in Engineering Design*, 25(4), 325–349. <https://doi.org/10.1007/s00163-014-0180-x>
- Shao, L., Migimatsu, T., Zhang, Q., Yang, K., & Bohg, J. (2021). Concept2Robot: Learning Manipulation Concepts from Instructions and Human Demonstrations. *The International Journal of Robotics Research*, 40(12–14), 1419–1434. <https://doi.org/10.1177/02783649211046285>
- Sheth, J. A. (2010). Questionnaire Design. In *Wiley international encyclopedia of marketing*. Wiley. <https://doi.org/10.1002/9781444316568>
- Shiralkar, K. (2017). Manufacturing Plant Layout Design – An Artificial Intelligence Approach. *International Journal for Research in Applied Science and Engineering Technology*, V(XI), 1791–1794. <https://doi.org/10.22214/ijraset.2017.11259>
- Siddharth, L., Blessing, L., & Luo, J. (2022). Natural Language Processing in-and-for Design Research. *Design Science*, 8. <https://doi.org/10.1017/dsj.2022.16>
- Singh, S. K., Rai, R., Khawale, R. P., Patel, D., Bielecki, D., Nguyen, R., Wang, J., & Zhang, Z. (2024). Deep Learning in Computational Design Synthesis: A Comprehensive Review. *Journal of Computing and Information Science in Engineering* (Vol. 24, Issue 4). American Society of Mechanical Engineers (ASME). <https://doi.org/10.1115/1.4064215>
- Sloterdijk, P., Stiegler, B., Lemmens, P., & Hui, Y. (2017). Reframing the

- Technosphere. *Journal for Contemporary Philosophy* (Issue 2).
- Song, B., Zurita, N. F. S., Nolte, H., Singh, H., Cagan, J., & McComb, C. (2021). When Faced with Increasing Complexity: The Effectiveness of Artificial Intelligence Assistance for Drone Design. *Journal of Mechanical Design*, 144(2). <https://doi.org/10.1115/1.4051871>
- Song, Y., & Luximon, Y. (2020). Trust in AI agent: A systematic review of facial anthropomorphic trustworthiness for social robot design. *Sensors* (Vol. 20, Issue 18, pp. 1–21). MDPI AG. <https://doi.org/10.3390/s20185087>
- Strobel, H., Kinley, J., Krueger, R. F., Beyer, J., Pfister, H., & Rushton, G. (2022). GenNI: Human-Ai Collaboration for Data-Backed Text Generation. *IEEE Transactions on Visualization and Computer Graphics*, 28(1), 1106–1116. <https://doi.org/10.1109/tvcg.2021.3114845>
- Taherdoost, H. (2022). Designing a Questionnaire for a Research Paper: A Comprehensive Guide to Design and Develop an Effective Questionnaire. *Asian Journal of Managerial Science*, 11(1), 8–16. <https://doi.org/10.51983/ajms-2022.11.1.3087>
- Tarkhova, L. (2023). Support for Decision-Making in Checking the Level of Quality of Student Research Works Based on Automated Text Analysis. *Culture and Education*, 35(4), 1068–1102. <https://doi.org/10.1080/11356405.2023.2255819>
- Testoni, A., & Bernardi, R. (2022). Garbage In, Flowers Out: Noisy Training Data Help Generative Models at Test Time. *Italian Journal of Computational Linguistics*, 8(1). <https://doi.org/10.4000/ijcol.974>
- Thißen, M., & Hergenröther, E. (2023). *Why Existing Multimodal Crowd Counting Datasets Can Lead to Unfulfilled Expectations in Real-World Applications*. <https://doi.org/10.24132/csrn.3301.5>
- Tremblay, M. C., Hevner, A. R., & Berndt, D. J. (2010). Focus Groups for Artifact Refinement and Evaluation in Design Research. *Communications of the Association for Information Systems*, 26. <https://doi.org/10.17705/1cais.02627>
- Tsang, Y. P., & Lee, C. K. M. (2022). Artificial intelligence in industrial design: A semi-automated literature survey. In *Engineering Applications of Artificial Intelligence* (Vol. 112). Elsevier Ltd. <https://doi.org/10.1016/j.engappai.2022.104884>
- Tubadji, A., Huang, H., & Webber, D. J. (2021). Cultural proximity bias in AI-acceptability: The importance of being human. *Technological Forecasting and Social Change*, 173. <https://doi.org/10.1016/j.techfore.2021.121100>
- Vadla, M. K. S. (2024). Enhancing Product Design Through AI-Driven Sentiment Analysis of Amazon Reviews Using BERT. *Algorithms*, 17(2), 59. <https://doi.org/10.3390/a17020059>
- Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and Design in the Age of Artificial Intelligence. *Journal of Product Innovation Management*, 37(3),

- 212–227. <https://doi.org/10.1111/jpim.12523>
- Vianello, A., Laine, S., & Tuomi, E. (2023). Improving Trustworthiness of AI Solutions: A Qualitative Approach to Support Ethically Grounded AI Design. *International Journal of Human-Computer Interaction*, 39(7), 1405–1422. <https://doi.org/10.1080/10447318.2022.2095478>
- Williams, D. D., South, J. B., Yanchar, S. C., Wilson, B., & Allen, S. (2011). How Do Instructional Designers Evaluate? A Qualitative Study of Evaluation in Practice. *Educational Technology Research and Development*, 59(6), 885–907. <https://doi.org/10.1007/s11423-011-9211-8>
- Vinnervik, P. (2022). Implementing programming in school mathematics and technology: teachers' intrinsic and extrinsic challenges. *International Journal of Technology and Design Education*, 32(1), 213–242. <https://doi.org/10.1007/s10798-020-09602-0>
- Wang, X., & Dong, A. (2008). A case study of computing appraisals in design text. In *Design Computing and Cognition'08: Proceedings of the Third International Conference on Design Computing and Cognition* (pp. 573–592). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-8728-8_30
- Wang, Y. Y., & Chuang, Y. W. (2024). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies*, 29(4), 4785–4808. <https://doi.org/10.1007/s10639-023-12015-w>
- Wu, J., Xing, B., Si, H., Dou, J., Wang, J., Zhu, Y., & Liu, X. (2020). Product Design Award Prediction Modeling: Design Visual Aesthetic Quality Assessment via DCNNs. *IEEE Access*, 8, 211028–211047. <https://doi.org/10.1109/access.2020.3039715>
- Wu, X. (2024). Factors Influencing Users' Satisfaction with Online Agricultural Products in China's First and Second-Tier Cities Under the Epidemic. *E-Bangi Journal of Social Science and Humanities*, 21(1). <https://doi.org/10.17576/ebangi.2024.2101.22>
- Xia, Y., Singhal, V. R., & Zhang, G. P. (2016). Product Design Awards and the Market Value of the Firm. *Production and Operations Management*, 25(6), 1038–1055. <https://doi.org/10.1111/poms.12525>
- Xu, L., Qin, S., Wang, P., & Gao, J. (2019). Research review on artificial intelligence technology to provide design of man-machine interaction in industry and product design. *Research Review on Artificial Intelligence Technology to Provide Research Review on Artificial Intelligence Technology* (pp. 12–20). <https://aisel.aisnet.org/iceb2019>
- Yan, J., & Li, B. (2020). Research hotspots and tendency of intelligent manufacturing. *Kexue Tongbao/Chinese Science Bulletin* (Vol. 65, Issue 8, pp. 684–694). Chinese Academy of Sciences. <https://doi.org/10.1360/N972019-00125>

- Yang, W., Su, J., Zhang, S., Qiu, K., & Zhang, X. (2021). Intelligent Design of Product Forms Based on Design Cognitive Dynamics and a Cobweb Structure. *Computational Intelligence and Neuroscience*, 2021, 1–17. <https://doi.org/10.1155/2021/6654717>
- Yang, Y. (1989). Survey Steered Design: Evaluating User Recovery and Command Reuse Support by Questionnaire. *Behaviour and Information Technology*, 8(6), 437–459. <https://doi.org/10.1080/01449298908914573>
- Yang, Z., Bao, D., & Shen, D. (2022). Exploring design students' cognition and application of different levels of Chinese cultural elements. *International Journal of Technology and Design Education*, 32(1), 645–665. <https://doi.org/10.1007/s10798-020-09606-w>
- Yang, Z., Chen, Y., & Liu, Z. (2018). Elliptic Fourier analysis and perceptual matching for the evaluation of bioinspired sketching in conceptual design. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, 32(1), 92–107. <https://doi.org/10.1017/S0890060417000257>
- Yao, C.-C. J. (2020). Research on Logistics Distribution Path Analysis Based on Artificial Intelligence Algorithms. *International Journal of Biometrics*, 12(1), 100. <https://doi.org/10.1504/ijbm.2020.10027203>
- Yoshida, K. (2023). Improving the Accuracy of Automated Gout Flare Ascertainment Using Natural Language Processing of Electronic Health Records and Linked Medicare Claims Data. *Pharmacoepidemiology and Drug Safety*, 33(1). <https://doi.org/10.1002/pds.5684>
- Yu, L., Sinha, A. A., Wang, M. M., Chen, H., Berg, T. L., & Zhang, N. (2022). CommerceMM: Large-Scale Commerce MultiModal Representation Learning With Omni Retrieval. <https://doi.org/10.48550/arxiv.2202.07247>
- Yüksel, N., Börklü, H. R., Sezer, H. K., & Canyurt, O. E. (2023). Review of artificial intelligence applications in engineering design perspective. In *Engineering Applications of Artificial Intelligence* (Vol. 118). Elsevier Ltd. <https://doi.org/10.1016/j.engappai.2022.105697>
- Zhang, F. (2022). Design and Implementation of Industrial Design and Transformation System Based on Artificial Intelligence Technology. *Mathematical Problems in Engineering*, 2022. <https://doi.org/10.1155/2022/9342691>
- Zhang, H., Bai, X., & Ma, Z. (2022). Consumer reactions to AI design: Exploring consumer willingness to pay for AI-designed products. *Psychology and Marketing*, 39(11), 2171–2183. <https://doi.org/10.1002/mar.21721>
- Zhang, Y. (2024). Unlocking the Potential of Artificial Intelligence in Fashion Design and E-Commerce Applications: The Case of Midjourney. *Journal of Theoretical and Applied Electronic Commerce Research*, 19(1), 654–670. <https://doi.org/10.3390/jtaer19010035>

- Zheng, J. (2024). Product Appearance Aesthetics in Industrial Design Based on Variational Onsager Neural Network Optimized with Osprey Optimization Algorithm. *Journal of Electrical Systems*, 20(3s), 2762–2773. <https://doi.org/10.52783/jes.3172>
- Zhong, Y., Qu, J., Zheng, K., Luximon, Y., & Luo, J. (2024, June). Improving product design efficiency through integrated AI tools: an empirical study. In *the 15th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences*. AHFE Open Access. <https://doi.org/10.54941/ahfe1004664>
- Zhou, A., Liu, H., Zhang, S., & Ouyang, J. (2021). Evaluation and Design Method for Product Form Aesthetics Based on Deep Learning. *IEEE Access*, 9, 108992–109003. <https://doi.org/10.1109/access.2021.3101619>
- Zhou, J., Li, P., Zhou, Y., Wang, B., Zang, J., & Meng, L. (2018). Toward New-Generation Intelligent Manufacturing. *Engineering* (Vol. 4, Issue 1, pp. 11–20). Elsevier Ltd. <https://doi.org/10.1016/j.eng.2018.01.002>
- Zhou, J., Xu, T., Ren, S., & Guo, K. (2020). Two-Stage Spatial Mapping for Multimodal Data Fusion in Mobile Crowd Sensing. *IEEE Access*, 8, 96727–96737. <https://doi.org/10.1109/access.2020.2995268>
- Zhou, M., Cheng, R., Lee, Y. J., & Yu, Z. (2018). A visual attention grounding neural model for multimodal machine translation. *arXiv preprint arXiv:1808.08266*. <https://doi.org/10.18653/v1/d18-1400>
- Zhu, H., Wang, Z., Shi, Y., Hua, Y., Xu, G., & Deng, L. (2020). Multimodal Fusion Method Based on Self-Attention Mechanism. *Wireless Communications and Mobile Computing*, 2020, 1–8. <https://doi.org/10.1155/2020/8843186>
- Zhu, J., Villareale, J., Javvaji, N., Risi, S., Löwe, M., Weigelt, R., & Hartevelde, C. (2021, May). Player-AI interaction: What neural network games reveal about AI as play. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-17). <https://doi.org/10.1145/3411764.3445307>
- Zhu, Z. (2023, May). Research on artificial intelligence product design method based on product semantics. *International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2022)* (Vol. 12604, pp. 578-583). SPIE. <https://doi.org/10.1117/12.2674615>
- Zong, L., & Wang, N. (2022). Research on the Decision Model of Product Design Based on a Deep Residual Network. *Scientific Programming*, 2022. <https://doi.org/10.1155/2022/8490683>

Appendix A. Ethical Approval



To Luximon Yan (School of Design)
From SIU Kin Wai Michael, Chair, Departmental Research Committee
Email sdmsiu@ Date 11-Feb-2020

Application for Ethical Review for Teaching/Research Involving Human Subjects

I write to inform you that approval has been given to your application for human subjects ethics review of the following project for a period from 10-Oct-2019 to 27-Aug-2022:

Project Title: Artificial intelligence enhanced design process
Department: School of Design
Principal Investigator: Luximon Yan
Project Start Date: 10-Oct-2019
Reference Number: HSEARS20191010006

You will be held responsible for the ethical approval granted for the project and the ethical conduct of the personnel involved in the project. In case the Co-PI, if any, has also obtained ethical approval for the project, the Co-PI will also assume the responsibility in respect of the ethical approval (in relation to the areas of expertise of respective Co-PI in accordance with the stipulations given by the approving authority).

You are responsible for informing the Human Subjects Ethics Sub-committee in advance of any changes in the proposal or procedures which may affect the validity of this ethical approval.

SIU Kin Wai Michael

Chair

Departmental Research Committee (on behalf of Human Subjects Ethics Sub-Committee)

Appendix B. Consent form in Study 1, 2, 4 & 5



CONSENT TO PARTICIPATE IN RESEARCH

ARTIFICIAL INTELLIGENCE ENHANCED DESIGN PROCESS

I _____ hereby consent to participate in the captioned research conducted by _____.

I understand that information obtained from this research may be used in future research and published. However, my right to privacy will be retained, i.e. my personal details will not be revealed.

The procedure as set out in the attached information sheet has been fully explained. I understand the benefit and risks involved. My participation in the project is voluntary.

I acknowledge that I have the right to question any part of the procedure and can withdraw at any time without penalty of any kind.

Name of participant _____

Signature of participant _____

Name of Parent or Guardian (if applicable) _____

Signature of Parent or Guardian (if applicable) _____

Name of researcher _____

Signature of researcher _____

Date _____

Hung Hom Kowloon Hong Kong 香港九龍紅磡
Tel 電話 (852) 2766 5111 Fax 傳真 (852) 2784 3374
Email 電郵 polyu@polyu.edu.hk
Website 網址 www.polyu.edu.hk