INTEGRATED SYSTEM WITH DIGITAL TWIN TECHNOLOGY FOR ENHANCING SAFETY IN CRANE LIFTING OPERATIONS

MUDASIR HUSSAIN

PhD

The Hong Kong Polytechnic University

2025

The Hong Kong Polytechnic University

Department of Civil and Environmental Engineering

Integrated System with Digital Twin Technology for Enhancing Safety in Crane Lifting Operations

Mudasir Hussain

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

Doctor of Philosophy

Jan 2025

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own 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.

Signature: _____

Name of Student: Mudasir Hussain

DEDICATION

To

My late father, Mr. Muhammad Zaman, and mother, Mrs. Bakhtawara,

My brothers and sisters,

My beloved wife, Shahida,

My dearest daughter, Hania Mudasir,

My mentors and teachers,

ABSTRACT

The construction industry drives economic growth in both developing and developed countries, leading to an increasing demand for construction projects and activities. Modern construction practices increasingly adopt off-site prefabrication of heavy modules due to limited onsite space, which requires transportation and crane-based installation. Traditionally, certified operators manage lifting operations based on manual judgment, which often leads to accidents caused by occlusion, collision, overloading, or unsafe practices. According to data from the Hong Kong Housing Authority, 38% of crane-related accidents occur during lifting operations, underscoring the urgent need for improved safety measures. In response, Hong Kong introduced the Smart Site Safety System (4S) in 2023. However, 4S relies on manual reporting and suffers from limitations in accuracy and adaptability, necessitating advanced technological solutions to enhance safety and reduce risks in lifting operations.

This thesis addresses three types of hazards during crane lifting operations and covers all contributing factors associated with these risks. These risks include occlusions/collisions, overloading, and unsafe practices. First, for occlusions/collisions, existing Computer-Aided Lift Path Planning and Replanning (CALPP-RP) are using parallel computing systems, but there are several challenges in parallel computing systems, including computational inefficiency, limited exploration capability, instability due to suboptimal problem division, and communication latency in dynamic environments. To overcome these limitations, this thesis proposes an integrated computing system for occlusion-free and collision-free path planning and replanning of robotized cranes. The workspace is modeled in Unity 3D and utilizes hybrid algorithms, combining A* for Configuration Space (c-space) exploration and Genetic Algorithms (GA) for path optimization. The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is employed to select optimal and sub-optimal paths, while raycasting is used for occlusion and collision detection. The system integrates a Decision Support System

(DSS) for proactive planning and a Path Re-Planner (PRP) for real-time updates. Pilot testing demonstrated improvements of 83.33% in computation speed and 33.32% in overall performance in high-dimensional environments compared to traditional sequential and parallel computing systems. CALPP-RP systems integrated with the proposed computing framework help avoid occlusion and collisions during lifting operations. Although cranes may operate for up to 60 years of Remaining Useful Life (RUL), they are subjected to cyclic and variable amplitude loads over time. These loads induce fatigue and stress in the crane structure, which accumulate and may eventually exceed critical thresholds, potentially leading to structural failure. Therefore, a system is needed to predict this degradation based on real-time sensor data.

Second, aging tower cranes face structural degradation, yet their lifting capacity over time remains underexplored. This thesis presents a Digital Twin-Driven (DTD) model to predict the degraded lifting capacity of aging cranes. The DTD model integrates fatigue analysis, real-time data, and Machine Learning (ML) models, achieving high accuracy (Mean Squared Error (MSE) = 0.2253, coefficient of determination (R^2) = 0.9973) in predicting degraded load charts over a 70-year lifespan. These predictions enhance safety monitoring, enabling operators to mitigate risks and prevent structural failures. Even after addressing occlusion, collision, and structural degradation, crane accidents occur due to unsafe practices. Therefore, there is a need for an automated safety risk assessment system that monitors all lifting activities and operations in real-time.

Third, traditional safety monitoring relies heavily on manual reporting, often error-prone and limiting its effectiveness. To advance 4S technologies, a cascaded model for automated safety risk assessment is proposed. It integrates Super-Resolution Generative Adversarial Networks (SRGAN) for image preprocessing, Real-Time Detection Transformer-Large (RT-DETR-L) for crane detection, self-DIstillation with NO labels (DINOv2) for safety classification, and Vision Transformer (ViT) for activity recognition. Comprehensive risk values are calculated using

probability matrices, triggering real-time warnings for high-risk scenarios. The model demonstrates superior performance, achieving 92.10% detection precision, 99.25% safety classification accuracy, and 99.47% activity classification accuracy, with an inference speed of 0.70 seconds, providing a reliable, real-time monitoring solution.

This research significantly advances crane safety in construction through integrated computing, predictive maintenance, and automated risk assessment. These innovations enhance efficiency, accuracy, and safety, offering scalable solutions to address critical challenges in dynamic construction environments.

Keywords: Integrated computing system, Digital twin, Safety risk assessment, Cranes, advanced computational technologies.

LIST OF PAPERS

The thesis is based on the following papers:

1) **M Hussain**, Z Ye, HL Chi*, SC Hsu (2024). 'Predicting degraded lifting capacity of aging tower cranes: A digital twin-driven approach,' *Advanced Engineering Informatics*, 59, p.102310.doi:10.1016/j.aei.2023.102310.
2) **M Hussain**, Z Ye, HL Chi*, SC Hsu (2025). 'Automated safety risk assessment for crane operations using cascade learning,' *Automation in Construction (Under review)*
3) **M Hussain**, Z Ye*, SC Hsu (2025). 'Integrated computing system for occlusion-free and collision-free path planning and replanning for robotized cranes in dynamic environments,' Automation in Construction (*Ready to Submit*)

Other papers of the author during the PhD program are not included in this thesis:

1) **M Hussain**, Z Ye, B Zheng, CM Lam, SC Hsu, TM Chan – Sustainable design of high-strength steel structural components: Process-based life cycle assessment with uncertainty and sensitivity analysis. *Journal of Constructional Steel Research* (*Under review*)
2) **M Hussain**, Z Ye, SC Hsu – Parametric BIM-based environmental and economic performance of high-strength steel grades in high-rise building designs. *Journal of Constructional Steel Research* (*Under review*)
3) **M. Hussain**, Shahida, Z Ye, SC Hsu – Valorization of egg white from discarded eggs as additives for sustainable mudbrick manufacturing: A comprehensive technical and environmental evaluation. *Construction and Building Materials* (*Under review*)
4) **M Hussain**, B Zheng, HL Chi, SC Hsu, JH Chen (2023). 'Automated and continuous BIM-based life cycle carbon assessment for Infrastructure Design Projects,' *Resources, Conservation and Recycling*, 190, p. 106848. doi:10.1016/j.resconrec.2022.106848.

5) B Zheng, **M Hussain**, Y Yang, APC Chan, HL Chi (2023). 'Trade-offs between accuracy and efficiency in BIM-LCA integration,' *Engineering Construction and Architectural Management*. doi:10.1108/ecam-03-2023-0270.
6) GA Anwar, **M Hussain**, MZ Akber, MA Khan, AA Khan (2023). 'Sustainability-oriented optimization and decision making of community buildings under seismic hazard,' *Sustainability*, 15(5), p. 4385. doi:10.3390/su15054385.
7) GA Anwar, MZ Akber, HA Ahmed, **M Hussain**, M Nawaz, J Anwar, WK Chan, HH Lee (2023). 'Sustainability-oriented optimization and decision making of community buildings under seismic hazard,' *Sustainability*, 15(5), p. 4385. doi:10.3390/su15054385.
8) I Mehmood, H Li, Y Qarout, W Umer, S Anwar, H Wu, **M Hussain**, MF Antwi-Afari (2023). 'Deep learning-based construction equipment operators' mental fatigue classification using wearable EEG Sensor Data,' *Advanced Engineering Informatics*, 56, p. 101978. doi:10.1016/j.aei.2023.101978.

Conference proceedings:

1) **M. Hussain**, HL Chi, B Zheng, SC Hsu – BIM-based life-cycle carbon assessment for sustainable tunnel design. *International Conference on Resource Sustainability (icRS) 2022*.
2) **M. Hussain**, Shahida, Z Ye, SC Hsu – Valorization of egg white from discarded eggs as additives for sustainable mudbrick manufacturing: A comprehensive technical and environmental evaluation. *International Conference on Resource Sustainability (icRS) 2024*.
3) **M Hussain**, Z Ye, SC Hsu – Parametric BIM-based environmental and economic performance of high-strength steel grades in high-rise building designs. *International Conference on Resource Sustainability (icRS) 2024*.

ACKNOWLEDGMENTS

Verily, all praise belongs to Almighty Allah, and salutations to Fourteen Infallible Masomeen (a.s), for granting me the courage and strength to complete my doctoral study.

I am profoundly grateful to my Chief Supervisor, Prof. Dr. Shu-Chien (Mark) Hsu, and Co-Supervisor, Prof. Dr. Hung-Lin Chi, whose exceptional mentorship has guided me from a novice to an independent researcher. Their unwavering support, insightful guidance, and holistic approach to my academic development have been instrumental to my success. Beyond research, their emphasis on physical well-being and healthy living has profoundly impacted both my academic and personal growth.

I sincerely thank The Hong Kong Polytechnic University for awarding me the prestigious PolyU Presidential PhD Fellowship Scheme (PPPFS) scholarship. The university's outstanding research facilities and environment have been crucial to my academic achievements. I also thank the General Office (GO) and Graduate School (GS) staff for their efficient and timely support.

I want to express my special thanks to my colleagues at PolyU: Dr. Vahid Asghari, Dr. Zhongnan Ye, Dr. Ziyue Yuan, Dr. Can Chen, Dr. Yunping Huang, Mr. Zixuan Kang, Ms. Shujie Xu, Ms. Xiaoyi Liu, Mr. Xinyu Yan, Ms. Rongyan Li, and Ms. Junyu Chen. Their scholarly discussions, critical feedback during group meetings, and camaraderie have significantly enriched my research experience.

I am deeply appreciative of my family: my brothers, Laiq, Sultan, Hamid, Zakir, Abid, Shakeel, and Adil; my sisters, Maeena, Tasleema, and Arifa; my father-in-law, Mr. Fida Hussain; my mother-in-law, Noor Nisa; and my beloved wife, Shahida, whose unwavering support and care during challenging times have been a constant source of strength. Your dedication, particularly

through your thoughtful encouragement and culinary expertise, has been invaluable throughout this journey.

Finally, I dedicate this achievement to my late father, Mr. Muhammad Zaman, and my mother, Mrs. Bakhtawara. Their unwavering love, sacrifices, and guidance have driven my accomplishments despite limited resources. My father's advice to pursue civil engineering, inspired by his work as a mason collaborating with engineers, profoundly influenced my journey.

Although I deeply miss my father, I pray that Allah grants him a high rank in Jannah. I hope he would be proud to see his son graduate from a top-ranked university with a PhD in Civil Engineering. To my beloved mother, I continue to pray for your health and well-being. Your support and inspiration have been the foundation of my success.

TABLE OF CONTENTS

THESIS CONTENTS

LIST OF TABLES

Table 5. 1: Overview of the selected activities in mobile and tower crane operations. 135

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Definition
ACID	Alberta Construction Image Dataset
ACO	Ant Colony Optimization
ADC	Analogue to Digital Converter
AI	Artificial Intelligence
AIFI	Attention-based Intra-scale Feature Interaction
AP	Average Precision
ATB	Anti-Two-Blocking
AWBiFPN	Augmented Weighted Bidirectional Feature Pyramid Network
BCE	Binary Cross-Entropy
Bi-FPN	Bi-directional Feature Pyramid Network
C	Constraints
CALPP-RP	Computer-Aided Lift Path Planning and Replanning
CCFM	CNN-based Cross-scale Feature-fusion Module
CD	Comprehensive Data
CNN	Convolutional Neural Networks
C-space	Configuration Space
CSPNet	Cross-Stage Partial Network
CV	Computer Vision
CV-CPM	Computer-Vision-Based Construction Progress Monitoring
DC	Direct Current
DD	Dynamic Data
DDM	Data-Driven Methods
DETR	DEtection TRansformer
DINOv2	self-DIstillation with NO labels
DL	Deep Learning
DOF	Degrees of Freedom
DRRT	Dynamic Rapidly Exploring Random Trees
DSS	Decision Support System
DTD	Digital Twin-Driven
ED	Environmental Data
ERRT	Extended Rapidly Exploring Random Trees

FEA	Finite Element Analysis
FHAP	Fuzzy Hierarchical Analysis Process
FNN	Feedforward Neural Networks
FOS	Factor of Safety
FPS	Frame Per Second
FTE	Full-Time Equipment
GA	Genetic Algorithm
GAN	Generative Adversarial Networks
GFLOPs	Giga Floating-Point Operations
GIoU	Generalized Intersection Over Union
GUI	Graphic User Interface
HD	Historical Data
HPC	High-Performance Computer
IoT	Internet of Things
IoU	Intersection over Union
IQR	Interquartile Range
LC	Lifting Capacity
LiDAR	Light Detection and Ranging
LMI	Load Moment Indicators
LSTM	Transformer Encoder-Long Short-Term Memory
mAP	Mean Average Precision
MBM	Model-Based Methods
MDM	Multilevel Depth Map
ML	Machine Learning
MS-COCO	Microsoft Common Objects in Context Dataset
MSE	Mean Square Error
NC	Without Constraints
NMS	Non-Maximum Suppression
PANet	Path Aggregation Network
PGA	Parallel Genetic Algorithm
PPE	Personnel Protection Equipment
PReLU	Parametric Rectified Linear Unit
PRP	Path Re-Planner

PSNR	Peak Signal-to-Noise Ratio
PSO	Particle Swarm Optimization
R^2	Coefficient of Determination
RAM	Risk Assessment Model
ReFuSeAct	Representation Fusion Using Self-supervised Learning for Activity Recognition
ReLU	Rectified Linear Unit
RF	Random Forest
RNN	Recurrent Neural Networks
RRT	Rapidly Exploring Random Trees
RSS	Resident Set Size
RT-DETR-L	Real-Time Detection Transformer-Large
RUL	Remaining Useful Life
SA	Simulated Annealing
SCP	Sequential Convex Programming
SD	Static Data
SGD	Stochastic Gradient Descent
Sp-D	Specification Data
SRGAN	Super-Resolution Generative Adversarial Networks
SSIM	Structural Similarity
SSSS	Smart Site Safety System
SVM	Support Vector Machine
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
UAV	Unmanned Aerial Vehicles
UBDA	University Big Data Analytics
UI	User Interface
VGG-19	Visual Geometry Group-19
ViT	Vision Transformer
VMS	Virtual Memory Size
YOLO	You-Only-Look-Once

*In alphabetical order

LIST OF NOTATIONS AND SYMBOLS

Notations and symbols	
ε	Obstacle penalty.
λ	Decay parameter.
$\left(X^i\right)^1$	Start configuration.
$\left(X^i\right)^1$	Start configuration of the obstacle. $X_j^i(t_i) = \left(X^i\right)^1 \forall\ (i,j)$.
$\left(X^i\right)^N$	Pseudo end configuration.
$\left(X^i\right)^{end}$	End configuration.
$\left(X^i\right)^{end}$	End configuration of the obstacle. $X_j^i(t_i + h) = \left(X^i\right)^{end} \in V^i, \forall\ (i,j)$.
$\left(\boldsymbol{E}^i\right)^1$	The first edge at which "Rotation" and "Hoisting" are randomly changed while other DOFs remain unchanged.
$\left(\boldsymbol{E}^i\right)^{N-1}$	"Rotation is aligned with the orientation of the target point, while other DOFs keep unchanged.
$\left(\boldsymbol{E}^i\right)^{N-2}$	The only change is "Hoisting," while other DOFs remain unchanged.
$\left(\boldsymbol{E}^i\right)^j$	The last edge at which all DOFs changed except for "Rotation," where 1< j < N-2
(h)	The estimated cost of the path from start configuration $\left(X^i\right)^1$ to end configuration $\left(X^i\right)^{end}$.
C^3	Cubic spline curve.
$C_k(t)$	Set of cost.
$D_j^i(t)$	The crane view vector at time *t*.
H^i	Time horizon for *ith* iteration between t_i and $t_i + h$.
$P^i(t)$	Obstacle predicted position at time $t \in H^i$.
$R_s(t)$	Set of rays where $s \in [0, N]$.
T^i	Transform matric of the crane animation space for H^i.
V^i	Set of priority queue at time $t_i + h$.
$\dot{a}(t)$	Tangent vector of the crane path at time *t*.
$a_j^i(t)$	3D position in animation $a_j^i \in \mathbb{A}^i$, at time $t \in H^i$.
(f)	Evaluation function of the A*.
(g)	Cost to reach end configuration $\left(X^i\right)^{end}$ from the start configuration $\left(X^i\right)^1$.

$\boldsymbol{E^i}$	An edge between two adjacent configurations.
$\widehat{\boldsymbol{S}}$	Set of optimal and suboptimal paths after TOPSIS, such as $\widehat{\boldsymbol{S}} = \{s_1, s_2\}$
$\boldsymbol{X^i}, \boldsymbol{C}$	Set of configurations in the c-space.
$\mathbb{A}^i$	Set of animation for horizon H^i.
δ_t	Time step.
E	Exponential decay function.
G	Gaussian decay.
h	User-defined duration of the time horizon H^i.
σ	Standard deviation.
$E(x, \lambda)$	Exponential decay, $e^{-x/\lambda}$.
$G(x, \sigma)$	Gaussian decay, $e^{-x^2/(2\sigma^2)}$.
M	Population size in the c-space.
N	Number of configurations in each path.
S	Set of paths generated by A*, such as $S = \{s_1, s_2, s_3 \ldots s_n\}$
s	Sling or path or chromosome.
$\boldsymbol{S}$	Set of optimal paths generated by GA, such as $\boldsymbol{S} = \{s_1, s_2, s_3 \ldots s_n\}$
$\boldsymbol{f}$, $\boldsymbol{f}(s)$	Fitness function for GA.

*In alphabetical order.

CHAPTER 01

INTRODUCTION

This chapter presents an overview of the research background, motivation, and problem statement, focusing on critical topics such as crane safety during lifting operations, path planning and replanning, degradation in lifting capacity, and safety risk assessment. It also outlines the research questions, objectives, and thesis structure. The introduction integrates and synthesizes findings from the author's published and submitted papers, which form the foundation of this thesis.

1.1 Background and Motivation

The construction industry is a cornerstone of global economic growth, driving infrastructure development and urbanization across diverse project types, from residential and commercial buildings to industrial complexes and civil engineering works (Wong et al., 2015; Wong et al., 2008). Central to these activities are tower cranes, which serve as indispensable equipment for lifting, transporting, and positioning heavy materials and prefabricated components in constrained and dynamic environments (Zayed & Hussein, 2024; Pan et al., 2020; Taghaddos et al., 2018). While modern construction practices, such as off-site prefabrication of heavy modules, have intensified reliance on cranes due to space limitations and the need for precision (Abdelmageed et al., 2020; Zhang et al., 2021; Zheng et al., 2020). Crane operations remain critical across all construction projects, whether traditional or modular (Kayhani et al., 2021; Kiss et al., 2020).

Crane operations are inherently complex due to dynamic movements within unstructured, three-dimensional (3D) environments. Unlike 2D environments, where movement is constrained to a planar surface, crane operations must navigate a vertical axis with constantly changing spatial configurations, limited visibility, and varying elevations. These 3D constraints

introduce a higher degree of difficulty in avoiding collisions, maintaining clear lines of sight, and issuing timely warnings. As a result, cranes are highly susceptible to occlusion, obstruction by built structures or temporary installations, and unpredictable interference, all of which increase the risk of accidents and complicate real-time tracking and control.

Certified operators are responsible for managing lifting operations, navigating complex site layouts to deliver loads from initial staging areas to final positions (Al-Hussein et al., 2006; Ji et al., 2018; Lin et al., 2014). However, the widespread practice of "blind lifting," which relies on human judgment for obstacle detection and path planning, introduces systemic inefficiencies and safety hazards common to all crane-reliant projects. The primary hazard, collision with obstacles or structural components, persists as a leading cause of crane-related accidents, resulting in injuries, fatalities, and costly delays (Tam et al., 2011; Zhou et al., 2018a, 2018b). An occlusion-free and collision-free path planning system would enable safe crane operations with minimal accident risks (Ajmal et al., 2016; Ali et al., 2005).

In addition to spatial complexity, aging tower cranes face structural fatigue, corrosion, and deterioration caused by high-cycle loads, environmental exposure, and inadequate maintenance. These factors gradually reduce lifting capacity (Das et al., 2018; Gu et al., 2021; Guo et al., 2021; Pal et al., 2018; Vukelic et al., 2019). Despite these challenges, many operators continue to rely on outdated manufacturer load charts, which are calibrated under ideal conditions and fail to reflect real-world deterioration, thereby increasing the likelihood of catastrophic failures (Guo et al., 2021; Kulka et al., 2016; Pal et al., 2018). Using outdated specifications increases the likelihood of structural failures or catastrophic collapses.

Compounding these technical issues are systemic human factors, including insufficient training, limited situational awareness, and noncompliance with safety protocols, all of which contribute to preventable accidents across the industry (Gürcanli et al., 2015; Jung et al., 2022;

Lee et al., 2020). Although regulatory frameworks are in place to enforce safe practices, their effectiveness is often compromised by human error and inconsistent adherence (Xiao & Kang, 2021; Zhang et al., 2022). Therefore, there is a need for automated safety risk assessment systems capable of monitoring all crane activities in real time (Jeelani et al., 2016; Rahim Abdul Hamid et al., 2019; Shafique et al., 2019).

This complex safety landscape, exacerbated by the challenges of operating in 3D, unstructured environments, underscores the urgent need for an integrated computing system capable of addressing multiple critical dimensions: generating occlusion-free and collision-free lifting paths, dynamically predicting the degraded capacity of aging equipment, and providing real-time safety risk assessments. Such a system, adaptable across modular, traditional, and hybrid projects, would mitigate risks through automated, data-driven decision-making, ensuring safer and more efficient crane operations industry-wide.

1.2 Crane safety in lifting operations

Crane accidents frequently occur during various phases, including layout planning, assembly, dismantling, erection, and lifting operations, with most accidents occurring during lifting operations. These accidents often stem from occlusions, collisions, overloading, and unsafe practices.

According to the U.S. Bureau of Labor Statistics, there were 5,486 fatal work injuries in 2022, representing a 5.7% increase from the previous year. The fatality rate rose to 3.7 per 100,000 full-time equivalent workers (FTW), with one worker dying every 96 minutes, compared to every 101 minutes in 2021 (Bureau of Labor Statistics, 2023). Among these fatalities, 170 were due to "struck-by" accidents, while 59 resulted from being caught in, compressed by, or crushed by equipment or objects, as defined by OSHA (Brown et al., 2021). Additionally, there were 79,660 nonfatal injuries, with 32.8% caused by contact with objects or equipment (e.g., loaders,

excavators, graders, cranes), 31% from falls, slips, or trips, 25.2% from overexertion or other bodily reactions, and 4.7% from transportation incidents (Brown et al., 2020). "Struck-by" machinery incidents continue to be a major contributor to workplace fatalities worldwide, with particularly high rates reported in Hong Kong (19.98%), Japan (6.54%), Singapore (5.40%), and Australia (3.20%) (Chiang et al., 2017).

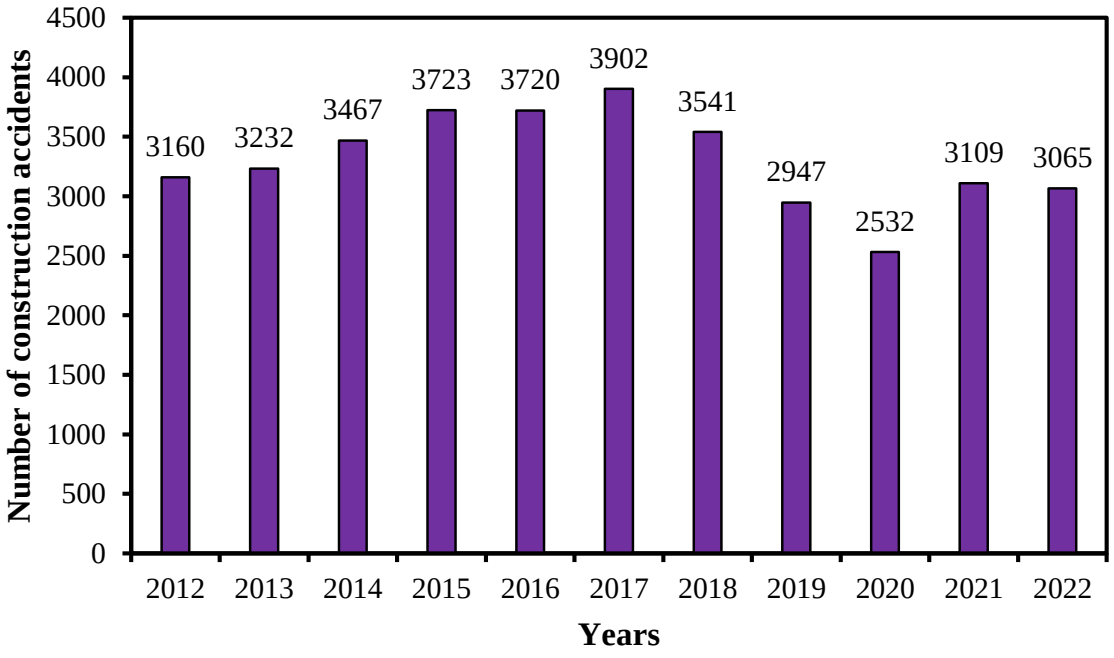
In Hong Kong alone, the Housing Authority recorded 3,109 construction accidents in 2021, an increase from 2,532 in 2020, as illustrated in Figure 1.1(a). The fatal work accident rates increased from 26.1 to 29.5 in 2022, as shown in Figure 1.1(b). A significant proportion of these accidents were linked to heavy objects or machinery, including cranes (ISSH, 2024). Notably, approximately 91.8% of construction workers operate near such objects or equipment, making them especially vulnerable to risks such as entrapment, compression, and being struck (Betit et al., 2022). As construction projects become larger and more complex, the reliance on heavy lifting equipment such as cranes has grown significantly (Hung et al., 2021a; Liu et al., 2021), elevated the associated risks (Shao et al., 2019).

Crane-related accidents in particular have shown a troubling upward trend. In 2022, 47 crane accidents were reported, up from 39 in 2021, as depicted in Figure 1.1(c). Of these, 38% occurred during lifting operations, 31% during climbing, assembly, or dismantling, and 23% were due to other factors, as shown in Figure 1.1(d). Overloading remains a major cause of lifting-related accidents, often due to the failure to account for the reduced structural capacity of aging cranes. These degradation-related failures have led to several catastrophic events. For instance, the 2022 collapse of a tower crane at the Sau Mau Ping construction site in Hong Kong resulted in three fatalities, as the operator failed to recognize the crane's deteriorated load capacity (Steven Chun-yin et al., 2022). Other notable incidents include the 2016 Dongguan crane collapse in China, which killed 18 people (Lan et al., 2017), and the 2008 New York City crane collapse, which caused seven deaths and multiple injuries (Linda Levine, 2008).

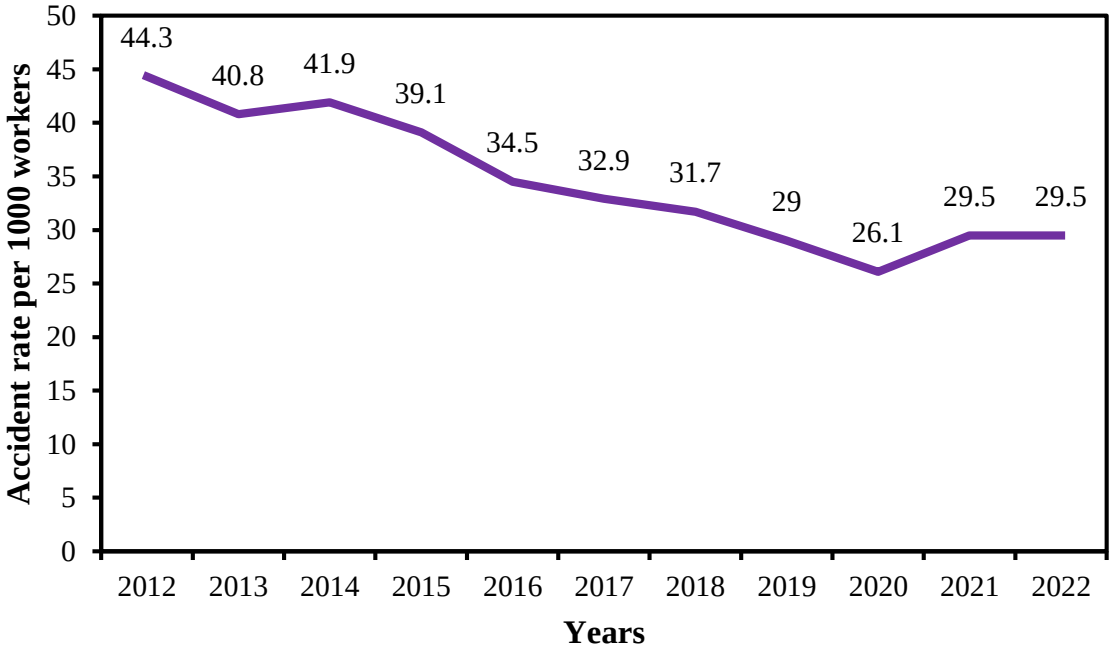
These events highlight the critical need for reliable detection of structural degradation in cranes and the prevention of overloading. Collisions and "struck-by" incidents emerge as leading causes of crane-related accidents, resulting in injuries, fatalities, project delays, and substantial financial losses.

Another significant contributing factor is unsafe crane operation. Accidents often result from improper usage, lack of training, low situational awareness, or operational fatigue (Gürcanli et al., 2015; Jung et al., 2022; Lee et al., 2020). Despite the existence of safety protocols, current safety enforcement measures frequently depend on manual supervision and subjective judgment (Xiao & Kang, 2021; Zhang et al., 2022). This can lead to delayed responses, missed warnings, and inconsistent implementation (Jeelani et al., 2016; Rahim Abdul Hamid et al., 2019; Shafique et al., 2019). The lack of systematic monitoring prevents timely intervention and hinders root cause analysis.

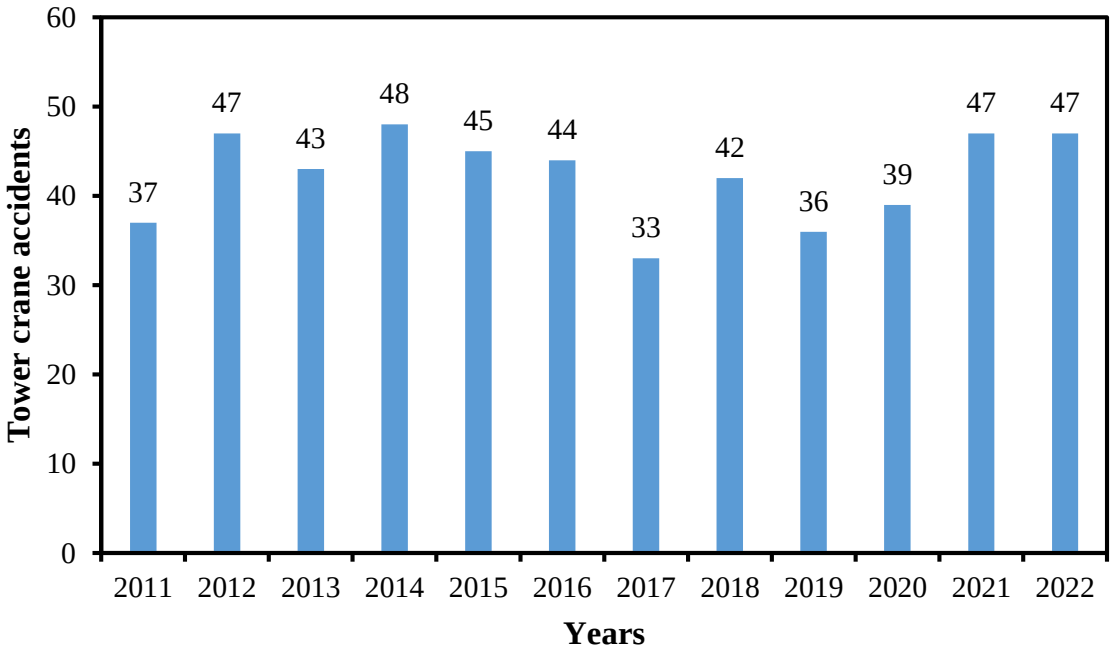
In summary, crane-related incidents; particularly those caused by overloading, structural degradation, collisions, and unsafe practices, represent a serious and escalating concern for construction safety worldwide. The growing scale of construction activities, increased deployment of heavy lifting machinery, and dependency on human-based safety monitoring collectively heighten the risk of severe accidents. These challenges underscore the urgent need for technological advancements and data-driven solutions to identify risks early and enhance on-site safety practices.



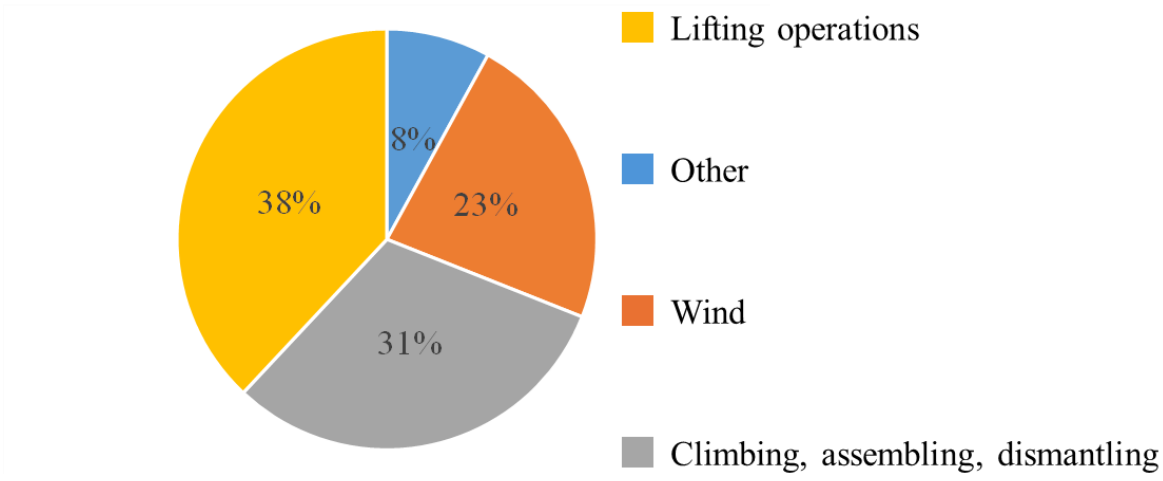
(a) Number of construction accidents per year in Hong Kong (Source: Hong Kong Housing Authority, www.housingauthority.gov.hk)



(b) Fatal accident rate per 1,000 workers in Hong Kong



(c) Number of tower crane accidents in Hong Kong (2011-2022)



(d) Tower crane accidents in various phases of crane operations

Figure 1. 1: Crane safety statistics during lifting operations (Source: Hong Kong Housing Authority).

1.3 Path planning and replanning

Efficient crane lift path planning and replanning are critical to ensuring optimal movement routes and minimizing operational risks. Traditionally, certified operators manage crane operations, lifting loads from loading areas to designated positions. However, these blind lifting

operations often lead to safety risks, project delays, injuries, fatalities, and financial losses. These challenges underscore the necessity for advanced computer-aided lift path planning and replanning (CALPP-RP) systems to enhance operational safety and efficiency. Such systems require rapid, real-time computation of collision-free lift paths to address the dynamic nature of construction environments.

CALPP-RP systems leverage advanced algorithms, intelligent decision-making frameworks, computer graphics, and simulations for real-time adaptation in construction projects (Chang et al., 2012a; Hu et al., 2021). They ensure safe and optimal crane lifting paths, enhancing safety and productivity (Ajeil et al., 2020). Despite these advancements, most existing systems treat planning and replanning as isolated processes, often yielding suboptimal results. In the dynamic environments of construction projects, continuous obstacle avoidance is paramount (Lei et al., 2013; Soltani et al., 2002). Planning algorithms must integrate real-time replanning steps to adjust paths based on environmental changes dynamically (Li et al., 2020; Zhang et al., 2012). Such adjustments require efficient algorithms that swiftly modify path segments based on new information (Dakulovi et al., 2011; Zhang et al., 2012). Consequently, developing an automated, real-time, integrated system for collision-free crane lift path planning and replanning is essential.

Various computational methods have been explored for real-time lift path planning and replanning, categorized broadly into deterministic, probabilistic, and evolutionary algorithms (Hu et al., 2021). Deterministic approaches, such as Dijkstra and Dynamic A* (D*), use predefined rules and heuristics to explore solution spaces and find optimal paths (Koenig et al., 2004; Soltani et al., 2002). These methods excel at recalibrating paths in response to environmental changes (Stentz, 1994). Extensions, such as heuristic-based D* (Stentz, 1995) and Two-Way D* (Dakulovi and Petrovi, 2011), reduce computation times by employing targeted strategies like heuristic focusing and occupancy grid maps. However, these algorithms

struggle with scalability and efficiency in high-dimensional configuration spaces (c-space), necessitating frequent recalculations and limiting adaptability.

Probabilistic methods, including Simulated Annealing (SA), Stochastic Gradient Descent (SGD), and Rapidly Exploring Random Trees (RRTs), utilize random processing to explore solution spaces (Hu et al., 2016; Zhang et al., 2012). While versatile, they face challenges in navigating high-dimensional c-spaces efficiently. For instance, Dynamic RRT (DRRT) improves computational efficiency by pruning only invalid segments of planned paths (Ferguson et al., 2014). Similarly, modified frameworks, such as Extended RRT (ERRT) and specialized DRRT implementations, balance path planning and execution in real-time (AlBahnassi and Hammad, 2012; Bruce and Veloso, 2014; Zhang and Hammad, 2012). Despite these advancements, achieving globally optimal solutions remains challenging due to biases introduced during sampling Bruce and Veloso (2018).

Evolutionary algorithms, such as Genetic Algorithm (GA), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO), simulate natural selection processes to refine potential solutions (Aslan et al., 2023; Tuncer and Yildirim, 2012). Hybrid approaches, such as PSO-SA combinations and multi-objective parallel genetic algorithms (Zhu et al., 2022), enhance computational efficiency in dynamic scenarios. However, challenges persist in balancing computational complexity, convergence speed, and solution accuracy. For example, Yang Wang (2007) hybrid GA improved obstacle avoidance but failed to achieve real-time adaptability. Similarly, GPU-based parallelization techniques enhance efficiency but struggle with scalability in complex environments (Dutta et al., 2020).

Despite these valuable contributions, several research gaps need to be addressed to improve the computational efficiency of tower crane lift path planning and replanning in dynamic environments: (1) delays arise from elevated computational intensity due to communication

overhead and load imbalances; (2) convergence rates are impacted by synchronization issues, load balancing, and algorithmic dependencies; (3) the success rate is hindered by a constrained exploration of the solution space; (4) instability is introduced through suboptimal subproblem division; (5) the system grapples with trade-offs between path smoothness and computational efficiency; and (6) challenges in communication latency and scalability emerge in intricate, high-dimensional, and dynamic environments.

An integrated computing system for occlusion-free and collision-free crane lift path planning and replanning in dynamic construction environments can address the limitations of sequential and parallel computing systems. With an automated, real-time occlusion-free and collision-free crane lift path planning and replanning system, external hazards can be effectively eliminated, enabling safer and more efficient crane operations. If an integrated computing system is incorporated into CALPP-RP systems, capable of avoiding occlusions and collisions in real-time, tower cranes could operate for up to 60 years of their remaining useful life (RUL). However, the most pressing concern then shifts to the degradation of lifting capacity in crane structures, driven by cyclic loading and adverse environmental conditions. It becomes imperative to accurately predict the degraded lifting capacity in aging cranes to ensure operational safety and structural integrity.

1.4 Degradation in lifting capacity during lifting operations

Tower cranes are indispensable for all types of construction projects, facilitating the lifting and precise placement of heavy materials, equipment, and tools (Code of Practice for Safe Use of Tower Cranes). Over time, these cranes are susceptible to structural fatigue, corrosion, and deterioration caused by high-cycle and variable-amplitude loads, harsh operating environments, and improper maintenance practices. Such factors contribute to the degradation of lifting capacity (LC), posing significant safety risks to workers and nearby properties (Das

et al., 2018; Gu et al., 2021; Kulka et al., 2018; Pal et al., 2018; Vukelic et al., 2019). Degraded LC increases the likelihood of accidents, resulting in severe injuries, fatalities, and financial losses (Lee et al., 2020).

Efforts to mitigate these risks have led to the adoption of safety monitoring systems in the construction industry, including load moment indicators (LMIs) (Fang and Cho, 2017; Kalairassan et al., 2017; Neitzel et al., 2001; Shaikh and Kumar, 2016), anti-two-blocking (ATB) systems (Rayco Wylie, 2012; Walbridge et al., 2022), and crane cameras (Chen et al., 2023; Elgendi et al., 2023). LMIs monitor crane loads and issue warnings when approaching or exceeding maximum LC thresholds (Fang and Cho, 2017). ATB systems prevent collisions by halting operations when the hook block nears the boom tip (Walbridge et al., 2022). Crane cameras provide visual monitoring to detect unsafe practices (Chen et al., 2023). While these systems enhance construction site safety, they fail to address LC degradation due to wear, fatigue, and environmental factors. Consequently, they cannot predict internal deterioration that could result in overloading incidents, as they lack mechanisms to account for long-term structural health.

Predicting LC degradation in aging cranes is further complicated by limited historical data, the inherent complexity of degradation mechanisms, and the unpredictability of external environmental factors. Research has explored two principal approaches to tackle these challenges: data-driven methods (DDM) and model-based methods (MBM). DDM leverages historical, real-time, and maintenance data to predict deterioration trends (Roman et al., 2019). While effective with abundant data, DDM can lack reliability without domain-specific knowledge of degradation mechanisms (Tran et al., 2012), limiting their applicability in practice (Roman et al., 2021). MBM, in contrast, models failure mechanisms such as cracks, wear, and stress to estimate performance degradation. However, these methods often require

extensive case-specific experiments and parameter identification, making them less efficient for complex systems like tower cranes.

Hybrid approaches integrating both methodologies have been proposed to overcome the limitations of DDM and MBM. For instance, Liu et al. (2021) combined multi-sensor data with interpretable MBM models, enhancing the accuracy and reliability of LC predictions. Although such methods improve safety monitoring, they add complexity and require the integration of diverse sensor data.

A promising hybrid approach involves using digital twin-driven (DTD) systems, which create virtual replicas of physical systems to enable real-time data acquisition, predictive modeling, and human-machine interaction (Jiang et al., 2022). DTD systems have been successfully applied in predictive maintenance, autonomous control, and performance evaluation across various domains, including manufacturing and construction (Cheng et al., 2020; Deebak and Al-Turjman, 2021; Jia et al., 2023; Kušić et al., 2023; J. Liu et al., 2021a, 2021b). Despite these advancements, limited research has applied DTD systems to predict LC degradation in aging tower cranes with real-time data, particularly in dynamic construction environments.

A framework using DTD systems can predict LC degradation over varying usage periods by combining theoretical and numerical fatigue and degradation analyses with real-time data collected from tower cranes. Machine learning models, such as Random Forest (RF) and Support Vector Machine (SVM), can be employed for feature selection and LC prediction. A scaled-down tower crane prototype can serve as a case study to validate this framework, with predictions of degraded load charts for usage periods ranging from 0 to 70 years. This approach enables continuous safety performance monitoring, addresses LC degradation in real-time, and significantly enhances safety on construction sites with aging cranes.

Considering that degraded LC presents critical safety risks, it is essential to combine it with an automated, occlusion-free, and collision-free lift path planning system. However, unsafe operations, near-miss incidents, and accidents remain pressing concerns. This thesis also highlights the need for an automated safety risk assessment framework to comprehensively mitigate these risks. Such a framework would systematically identify, evaluate, and control risks associated with lifting operations, thereby advancing safety and operational efficiency in construction environments.

1.5 Safety risk assessment in lifting operations

Construction sites vary in size, layout, and working conditions; however, they generally include workers, equipment, and materials (Shen et al., 2016; Teizer and Cheng, 2015). Heavy machinery and tools are essential for critical and complex activities to meet schedules (Zhou et al., 2018), but they pose risks if failures occur (Duarte et al., 2021; Shin et al., 2024). Accidents often stem from unsafe operation, improper use, lack of awareness, education, or skills (Gürcanli et al., 2015; Jung et al., 2022; Lee et al., 2020). Comprehensive monitoring and assessment of machine operations are crucial for understanding the causes of accidents and preventing them (Xiao and Kang, 2021; Zhang and Ge, 2022). Despite safety regulations and training, accidents continue to occur, leading to severe injuries, fatalities, and financial losses. Current safety measures often rely on manual judgments, which can be error-prone, delayed, or inconsistent (Hinze et al., 2016; Jeelani et al., 2016; Rahim Abdul Hamid et al., 2019; Shafique and Rafiq, 2019). Therefore, automated, advanced, accurate, reliable, and adaptive safety procedures are needed to identify, assess, evaluate, and monitor operational activities in real-time.

The Bureau of Labor Statistics reported that the United States recorded 5,486 fatal work injuries in 2022, reflecting a 5.70% increment from 2021. The fatal work injury rate rose to

3.70 per 100,000 full-time equivalent (FTE) workers, with worker fatality occurring roughly every 96 minutes, compared to every 101 minutes in 2021 (Bureau of Labor Statistics, 2023). Among these fatalities, 170 were attributed to being struck-by machines and 59 were due to incidents involving being caught-in, trapped-in, compressed-by, or crushed-by machines (Brown et al., 2021). Additionally, 79,660 nonfatal work-related injuries have been reported, with a significant portion resulting from contact with machines, falls, slips, trips, and overexertion (Brown et al., 2020). Machine-related struck-by incidents remain a leading cause of construction injuries globally (Hinze et al., 2005), with high fatality rates in Hong Kong (19.98%), Japan (6.54%), Singapore (5.40%), and Australia (3.20%) (Chiang et al., 2017). Hong Kong reported 3,065 construction accidents in 2022, with slips, falls, and incidents involving machinery as significant contributors (ISSH, 2024). Approximately 91.80% of construction workers operate near heavy machinery, making them vulnerable to hazards such as being caught-in, trapped-in, compressed-by, crushed-by, and struck-by equipment (Betit et al., 2022). Among these incidents, there were 47 crane accidents in 2022, up from 39 in 2021. Of these, 38.00% occurred during operation, 31.00% during climbing, assembling, or dismantling, and 23.00% were due to other factors. The critical importance of crane safety risk assessment during operation underscores the need for Smart Site Safety Systems (4S) to review and enhance current safety measures in the construction industry (Gu et al., 2022; Niu et al., 2016).

4S represents a construction site safety management proposal integrating computer vision and deep learning technologies (Awolusi et al., 2018; Johnson, 2002; Mustafa et al., 2023). This innovative system operates through three key elements: smart monitoring devices that capture and identify potential risks, networked data transmission systems, and centralized platforms that process visual information to coordinate responsive actions (Fang et al., 2020; Paneru and Jeelani, 2021; Seo et al., 2015; Xu et al., 2021). While traditional safety protocols rely heavily

on pre-construction risk assessments, 4S enables real-time hazard management during the active construction phases. This technological advancement significantly shifts from conventional practices by providing immediate detection and response capabilities for operational activities (Fung et al., 2012). The system leverages advanced computer vision algorithms to perform multiple critical functions simultaneously: detecting potential hazards, tracking movement patterns, recognizing unsafe behaviors, and classifying risk levels in real-time (Hung and Su, 2021; Khallaf and Khallaf, 2021). This comprehensive approach is particularly vital, given the dynamic and complex nature of modern construction environments (Sabuhi et al., 2021; Singh et al., 2024; Tylman et al., 2010).

As construction projects become increasingly sophisticated and more extensive in scale, the deployment of heavy machinery has grown proportionally to meet the productivity demands (Hung and Su, 2021a; Liu et al., 2021). This intensified machinery usage inherently increases operational risks. Since accidents involving heavy construction equipment typically result in more severe consequences than other construction incidents (Shao et al., 2019), implementing robust safety measures is paramount. Current research on computer vision-based safety systems primarily focuses on specific risk factors, such as collision prevention and unsafe operational practices (Fang et al., 2018). However, a truly effective safety management system must go beyond mere identification to incorporate a thorough risk assessment and informed decision-making processes. The existing research landscape shows a notable imbalance; while considerable attention is given to hazard identification and classification, there remains a significant gap in developing quantitative operational risk assessment methodologies (Fang et al., 2019; Schieg, 2006). This limitation potentially compromises the effectiveness of construction site safety measures and risk management strategies.

Recent deep learning models based on Generative Adversarial Networks (GANs), Recurrent Neural Networks (RNNs), and convolutional neural networks (CNNs) have demonstrated

reliable performance in construction applications, particularly in detecting predefined entities in video footage in near real-time. While researchers have attempted to enhance or modify specific components of GANs, RNNs, and CNNs, the overall network structures remain unchanged. A significant limitation of current safety monitoring approaches lies in their reactive nature, as they primarily focus on detecting unsafe behaviors only after they become apparent risks (Xiao et al., 2021). This retrospective approach overlooks the numerous latent risk factors at construction sites, creating potential blind spots in safety management. The practice of triggering alerts only when hazardous activities reach critical levels is fundamentally flawed from a preventive perspective (Fang et al., 2019; Xiao et al., 2021). A proactive and holistic approach to safety risk assessment is necessary, encompassing visible hazards and potential risks still in their dormant state. Such comprehensive oversight would enable construction managers to strategically identify areas requiring enhanced safety supervision and optimize resource allocation for preventive measures (Schieg, 2006). Moreover, this forward-thinking approach would strengthen crisis response capabilities, allowing safety teams to intervene before situations escalate to accidents. By identifying and addressing risk factors in their early stages, construction sites can either entirely prevent incidents or substantially reduce their impact when they do occur. For example, Fang et al. (2018) employed an Improved Faster R-CNN (IFaster R-CNN) to detect workers and equipment more accurately than other advanced detection methods. Despite these achievements, computer vision and deep learning face several challenges, including the requirement of large-annotated datasets, manual or semi-automated feature extraction, high computational demands, and issues of robustness and adaptability (Brozovsky et al., 2024; Dobrucali et al., 2024; Pandey et al., 2024). Researchers have developed specialized datasets to address some of these challenges and employed Transformer Networks for automated image-feature extraction. Notable examples include the work of Tajeen and Zhu, (2014), Lee et al.

(2022), and Xiao and Kang, (2020), who developed annotated datasets for various construction machines. Gao et al. (2024) utilized an augmented weighted bidirectional feature pyramid network to detect marine objects.

Despite recent advancements, the development of computer vision models tailored to diverse construction environments remains an ongoing challenge, particularly in terms of efficiency, reliability, accuracy, and adaptability (Al-Faris et al., 2020; Fang et al., 2023). Researchers are actively investigating the application of transformers and cascaded learning frameworks to improve the precision, robustness, and scalability of object-detection and classification systems (Martin et al., 2023). For instance, Kim et al., (2024) integrated various models into a cascaded framework to enhance the identification and classification of unsafe actions involving equipment in low-quality images, resulting in improved accuracy. Hou et al. (2023) utilized computer vision-based systems to compute safety risks and visualize construction sites. However, comprehensive safety risk assessments require more than just detection and computation systems for unsafe operations; they also require risk identification, evaluation, control, mitigation, monitoring, and documentation. Research gaps persist in this area due to the absence of automated systems for such integrations. Current computer vision models also face challenges related to high computational demands, inconsistent reliability, limited accuracy, and difficulties adapting to diverse construction settings. Furthermore, incorporating environmental conditions and varying equipment types to enhance model robustness presents a significant challenge. The labor-intensive nature of construction work and reliance on manual judgment add further complexity to integrating these advanced technologies. Therefore, a cascaded learning method was employed to automate the safety risk assessment in crane lifting operations.

The risks associated with lifting operations, including occlusion, collision, structural degradation, and unsafe practices, highlight the critical need for a comprehensive study to address these research challenges effectively.

1.6 Research objectives

This thesis aims to develop an integrated system for occlusion-free and collision-free crane lift path planning and replanning in dynamic environments. Additionally, it seeks to predict the degraded lifting capacity of aging tower cranes and implement automated safety risk assessments during the lifting operation phase. These assessments focus on identifying safe operations, unsafe operations, near-miss incidents, and actual incidents to enhance safety and operational efficiency.

Objective 01: An integrated computing system for occlusion-free and collision-free path planning and replanning for robotized cranes in dynamic environments.

This study developed an integrated computing system to facilitate safe crane lift operations by eliminating occlusions and collisions in dynamic construction environments. The workspace was modeled in Unity 3D and transformed into a configuration space (c-space) using robotic motion planning principles. A hybrid algorithm combining A* and Genetic Algorithm (GA) was implemented, where A* explored the c-space using evaluation functions and bending suppression, and GA optimized paths via selection, crossover, mutation, insertion, and deletion operators. The TOPSIS method was applied to identify both optimal and suboptimal paths. Raycasting techniques were used to detect potential occlusions and collisions in real-time. The system comprises a Decision Support System (DSS) for initial path planning using proximity and threshold distance data, and a Path Re-planner (PRP) to update local paths dynamically during operations. A pilot study in Unity validated the system's effectiveness.

Contribution: This objective contributes a novel real-time hybrid planning and replanning system that significantly enhances operational safety and efficiency in complex and dynamic crane workspaces.

Objective 02: Predicting degraded lifting capacity of aging tower cranes: A digital twin-driven approach.

This study proposed a Digital Twin-Driven (DTD) framework to predict the degradation of lifting capacity (LC) in aging tower cranes, which is critical for long-term safety and structural reliability. The framework integrates theoretical and numerical analyses of fatigue and structural deterioration with real-time data obtained from a scaled-down tower crane prototype. Machine learning models, such as Random Forest (RF) and Support Vector Machine (SVM), were employed for feature selection and LC prediction. The model generated degraded load charts for different usage periods ranging from 0 to 70 years, supporting proactive maintenance and decision-making.

Contribution: This objective introduces a predictive framework that combines digital twins and machine learning to provide real-time safety monitoring and lifespan forecasting for aging crane structures.

Objective 03: Automated safety risk assessment for crane operations using cascade learning.

This research presents a cascaded deep learning model for automated safety risk assessment in crane operations. The system incorporates a Super-Resolution Generative Adversarial Network (SRGAN) for image enhancement, RT-DETR-L for real-time crane detection, DINOv2 for safety status classification, and Vision Transformer (ViT) for activity recognition. Risk values are quantified through confidence-based probability matrices, and severity-likelihood analysis is used to determine impact scores. When the threshold value of 0.52 is exceeded, real-time

warnings are triggered for operators and site managers. Continuous monitoring via video feeds allows for systematic documentation and safety enforcement.

Contribution: This objective delivers a comprehensive, automated, and real-time safety risk assessment system that significantly improves the detection and mitigation of unsafe crane operations.

1.7 Thesis layout

The thesis is organized into six comprehensive chapters, each dedicated to a specific aspect of the study, as shown in Figure 1.2.

Chapter 01: Introduction provides the foundation for the thesis. It presents an overview of the research background, motivation, and problem statement, focusing on critical topics such as crane safety during lifting operations, path planning and replanning, degradation in lifting capacity, and safety risk assessment. It also outlines the research questions, objectives, and thesis structure.

Chapter 02: The literature review systematically examines various computing systems employed for path planning and replanning in dynamic environments, encompassing sequential, parallel, and integrated approaches. It further explores methodologies for predicting machine degradation, including analytical techniques, data-driven methods, and digital twin-based frameworks. Additionally, it discusses the application of computer vision technologies in construction safety management, emphasizing their role in risk assessment, object detection and tracking, activity recognition, classification, and analysis. The review also highlights the limitations of current studies, providing a foundation for addressing gaps in the existing body of knowledge.

Chapter 03: This thesis presents an integrated computing system for occlusion-free and collision-free path planning and replanning for robotized cranes operating in dynamic environments. The methodology is outlined, covering problem formulation for optimization, the conversion of workspace to configuration space (c-space), the sling formation process, and the implementation of a hybrid algorithm combining A* and Genetic Algorithm (GA). Additionally, the design of the decision-making system, collision detection through raycasting, and the development of a Decision Support System (DSS) and Path Re-planner (PRP) are discussed. A case study demonstrating the application of the integrated system in Unity 3D is included, where the system's performance is evaluated across various parameters. The study also compares sequential, parallel, and integrated computing systems, providing a comprehensive analysis of their effectiveness in dynamic crane operations.

Chapter 04: This chapter introduces a digital twin-driven (DTD) framework for predicting the degraded lifting capacity of aging tower cranes. It outlines a comprehensive approach, including the theoretical and numerical analysis workflows for assessing fatigue accumulation and degradation in lifting capacity. The chapter details the development process for constructing a DTD model tailored to tower crane degradation prediction. A case study involving a physical prototype of a tower crane is conducted to evaluate the model's accuracy, efficiency, and practical applicability. The findings are critically analyzed, offering insights into the model's performance and its potential contributions to enhancing the monitoring and maintenance of aging crane systems.

Chapter 05: This chapter presents an automated safety risk assessment framework for crane operations utilizing a cascade learning approach. The methodology encompasses data collection, preparation, and processing, including image preprocessing, risk identification, assessment, evaluation, mitigation, monitoring, and documentation. The architecture of the cascade learning pipeline is detailed, integrating SRGAN, RT-DETR-L, DINOv2, and ViT to

achieve robust performance. Results for each model within the pipeline are presented, along with an evaluation of the overall performance of the cascaded model in detection and classification tasks. The study analyzes performance metrics such as computational complexity, accuracy, reliability, and adaptability, benchmarking the proposed method against other state-of-the-art models to demonstrate its efficacy and potential for enhancing safety management in crane operations.

Chapter 06: The chapter on Conclusions, Limitations, and Future Work provides a comprehensive synthesis of the thesis, summarizing the key research findings, contributions, and implications. It highlights the study's impact while identifying areas where further research could extend the scope and applicability of the proposed models.

This chapter emphasizes the significance of the research outcomes and offers insights into potential directions for advancing the methodologies and applications explored. Each preceding chapter delivers a focused yet thorough examination of its respective research objectives, collectively contributing to the overarching aim of improving efficiency and safety in firefighting operations during building fire emergencies. The organizational structure of the thesis, which ensures coherence and progression, is depicted in Figure 1.2.

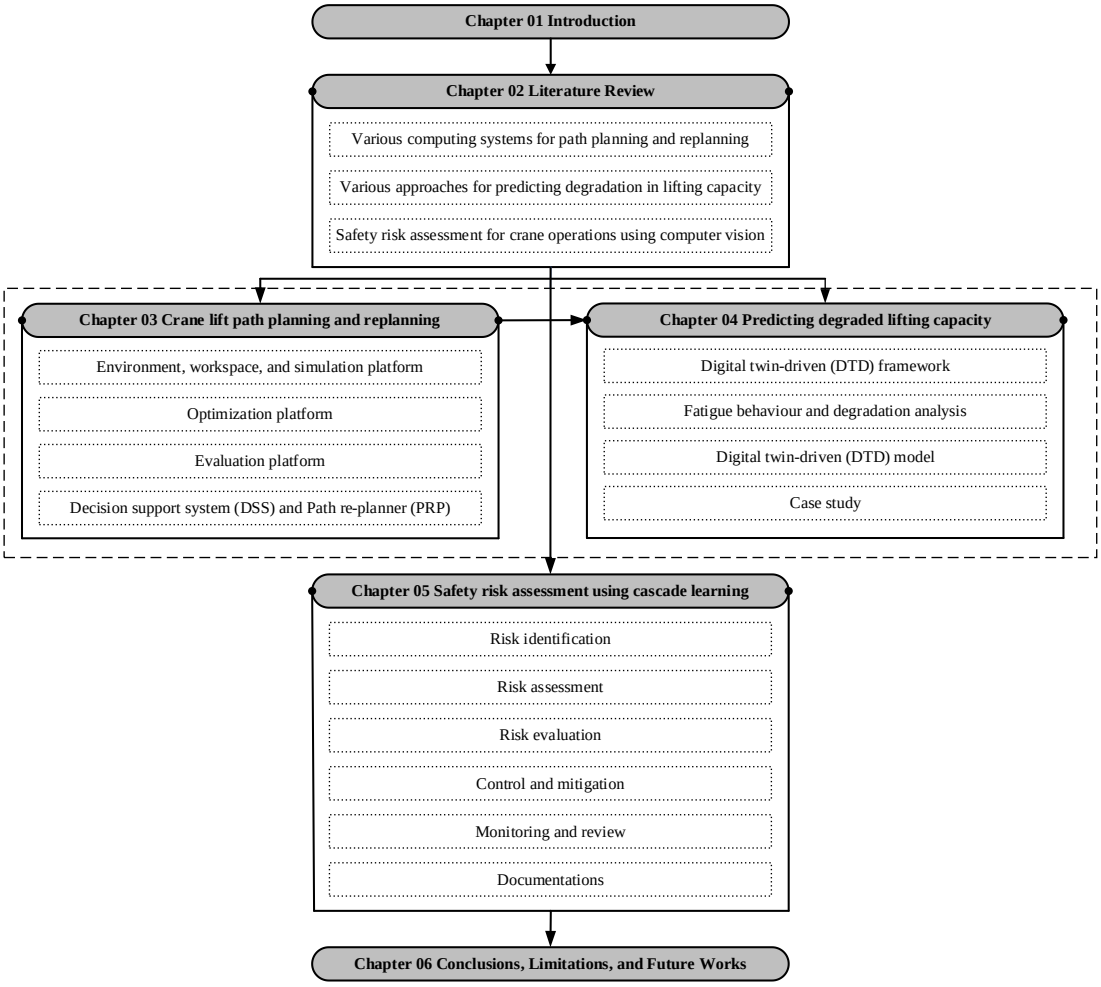


Figure 1. 2: Organization of the chapters in this thesis.

CHAPTER 02

LITERATURE REVIEW

The literature review provides a structured analysis of key research areas relevant to this study. It begins by examining the necessity of path planning and replanning during crane lifting operations (Section 2.1), followed by an exploration of computing systems used to enhance crane safety and efficiency in dynamic environments (Section 2.1.1). This includes an in-depth discussion of sequential (Section 2.1.2), parallel (Section 2.1.3), and integrated computing approaches (Section 2.1.4). The review then investigates methods for predicting machine degradation (Section 2.2), covering analytical models (Section 2.2.1), data-driven techniques (Section 2.2.2), and digital twin-based frameworks (Section 2.2.3). Furthermore, it addresses the importance of safety risk assessment (Section 2.3), with a particular focus on the application of computer vision technologies in construction safety (Section 2.3.1). This includes their use in risk assessment (Section 2.3.2), object detection and tracking (Section 2.3.3), as well as activity recognition, classification, and analysis (Section 2.3.4). The chapter also identifies limitations in the existing literature and outlines research gaps that warrant further investigation (Section 2.3.5). Finally, a summary of the entire chapter is presented (Section 2.4).

2.1 Crane lift path planning and replanning in dynamic environments

2.1.1 Crane safety in lifting operations

Cranes are indispensable on construction sites, valued for their efficiency in hoisting volumetric modules from loading areas to designated positions (Ali et al., 2024; Hussain et al., 2024). However, their operation presents significant safety and stability challenges, especially in dynamic environments characterized by moving objects, fluctuating wind conditions, and complex site layouts (Ali et al., 2024; Wu et al., 2024). Construction sites are inherently hazardous due to the simultaneous activities involving heavy machinery, mobile cranes,

laborers, and material handling (Sadeghi et al., 2021), all of which amplify the risk of accidents (Jiao et al., 2024; Sadeghi and Zhang, 2024). Cranes, as a central piece of equipment, are particularly associated with high safety risks (Jiang and Ding, 2024; Li, 2024), and numerous accidents worldwide have been attributed to human error, insufficient planning, equipment malfunction, and sudden environmental changes (Raj and Teizer, 2024). To address these issues, researchers have investigated the causes of injuries and fatalities across various crane-related activities such as layout planning (Sulankivi et al., 2014), installation and dismantling (Shin 2015), erection (Chang et al., 2012), and general operations (Sadeghi et al., 2023), identifying key risk factors including project conditions, environmental variables, human factors, and safety management practices (Shapira et al., 2014). Additionally, crane-related failures often stem from structural fatigue, wear and tear, stress, cracks, corrosion, and overall deterioration due to cyclic loading, harsh working environments, and inadequate maintenance. These factors contribute to the degradation of lifting capacity (LC) over time, underscoring the need for systematic research to quantify aging effects and develop strategies to monitor and mitigate such risks (Lee et al., 2020).

Moreover, tracking and predicting crane operations within unstructured, three-dimensional, and dynamic construction environments remains a substantial challenge. These environments are typified by irregular geometries, continuously evolving workspaces, and the presence of unpredictable obstacles—factors that complicate the real-time localization, path prediction, and motion control of cranes. The lack of fixed infrastructure, constantly changing layouts, and partial observability of the surroundings further exacerbate the difficulty in achieving accurate and reliable operation tracking. Consequently, advanced sensing, computing, and algorithmic strategies are required to overcome these limitations and ensure operational safety.

To ensure safe and uninterrupted operations, cranes must follow meticulously planned paths that avoid obstacles and prevent collisions. This requires computationally intensive optimization, taking into account crane capacity, load weight, and environmental conditions (Hussain et al., 2024). Real-time path planning and replanning have become critical for adapting to site dynamics, with researchers exploring deterministic, probabilistic, and evolutionary algorithms (Hu et al., 2021; Weingartshofer et al., 2023). Although achieving consistent improvements remains challenging, hybrid algorithms have demonstrated promise in addressing real-time operational demands (Roman et al., 2021). In response, advanced computing systems are being developed to enhance safety, efficiency, and adaptability in crane operations, potentially transforming safety protocols and management practices in construction environments (Bertram et al., 2006).

2.1.1.1 Sensor-based tracking and 4D BIM integration for crane operations

Recent advancements in construction automation have emphasized the significance of sensor-based tracking systems for crane operations, particularly in dynamic and spatially complex environments. Accurate and real-time tracking of crane components, including the boom, hook, and load, is essential for enhancing operational safety and coordination (Gheisari and Irizarry, 2016). Technologies such as global positioning systems (GPS), ultra-wideband (UWB), inertial measurement units (IMUs), and vision-based systems have been employed to monitor crane movements with varying degrees of spatial and temporal resolution (Li et al., 2018; Chen et al., 2020). These sensor-based methods, when integrated with 4D Building Information Modeling (BIM), offer the potential to visualize crane operations in conjunction with evolving construction schedules, enabling better planning and collision avoidance (Kim et al., 2013; Zhang et al., 2020). 4D BIM allows for the temporal linking of schedule data with 3D spatial models, which facilitates simulation-based decision-making for crane paths, lifting sequences, and resource allocation. Moreover, the fusion of real-time crane telemetry data with 4D BIM

platforms can support dynamic replanning and real-time progress monitoring (Ma et al., 2021). Another emerging focus is the verification of object constraints in 3D environments, which is critical for assessing spatial feasibility and safety during lifting tasks. Constraint-based modeling approaches have been developed to evaluate crane lift capabilities within confined geometries, taking into account obstructions, workspace limitations, and kinematic boundaries (Sacks et al., 2015; Li and Becerik-Gerber, 2011). Such frameworks support proactive hazard identification and constraint validation, minimizing the likelihood of unsafe or infeasible operations. Despite these advances, the integration of these technologies into a cohesive real-time decision-support system remains limited, highlighting the need for further research into scalable, interoperable platforms that combine sensing, modeling, and analytics for crane operation tracking and planning in construction environments.

2.1.2 Sequential computing systems for path planning and replanning

Sequential computing systems execute tasks one at a time, following a linear progression (Wan et al., 2016). This step-by-step processing facilitates understanding and logic tracing (Chen et al., 2014), proving effective for more straightforward path-planning problems in static environments. Researchers have applied sequential computing to enhance computational efficiency in vehicles, Unmanned Aerial Vehicles (UAVs), and robot path planning.

For vehicles, Bansal et al. (2016) introduced sequential path planning, demonstrating linear scalability for multi-vehicle planning. Zuo et al. (2019) approached path planning as determining a sequence of network nodes, using a sequence-to-sequence model to learn implicit forwarding paths from real-world data. In drone applications, Shi et al. (2021) employed sampling-based sequential optimization to calculate inspection paths. Aslan et al. (2023) adopted sequential computing with various algorithms to find optimal paths in dynamic 3D environments. For robotics, Chen et al. (2015) proposed a multi-agent path planning

algorithm based on sequential convex programming (SCP), gradually tightening collision constraints. Touzani et al. (2021) presented a sequential methodology to minimize time and movement durations. Fareh et al. (2020) achieved significant reductions in computational time (97.05%) and improved path quality (16.21%) for real-time algorithms by focusing on critical areas. Goutham et al. (2023) developed an efficient method for determining optimal paths through waypoint sequences.

However, sequential computing systems may encounter challenges with large-scale problems or complex algorithms due to their linear nature, potentially resulting in slower computation times. Limited parallelism can decrease performance for intensive tasks, and these systems may not fully exploit the capabilities of modern multi-core processors.

2.1.3 Parallel computing systems for path planning and replanning

Parallel computing systems execute multiple instructions or processes simultaneously, enhancing path planning tasks (Henrich, 1996). This approach divides problems into smaller, concurrent subproblems, improving speed and efficiency (Cai et al., 2016; Henrich, 1997). Multiple processors benefit algorithms divisible into independent subproblems (Tsai et al., 2011). This concept applies to transportation networks, robotics, UAVs, Light Detection and Ranging (LiDAR), and cranes.

In transportation, Liu et al. (2019) utilized parallel computing for efficiency in static and dynamic environments. For robotics, Weingartshofer et al. (2023) proposed a framework for industrial processes, considering all degrees of freedom (DOF) and tolerance. Gul et al. (2023) introduced a parallel strategy to enhance multi-robot space exploration while minimizing complexity. For UAVs, Cekmez et al. (2014) implemented a parallelized Ant Colony Optimization (ACO) algorithm on CUDA for path planning. Roberge et al. (2013) compared PGA and PSO, concluding PGA produced better trajectories. In LiDAR applications, Vo et al.

(2021) integrated a GA within a dual parallel framework to optimize flight paths. For cranes, Cai et al. (2018) used a prioritized multi-objective PGA for automatic path planning in dynamic environments. Dutta et al. (2020) employed a GPU-based parallelization for path replanning, leveraging an MDM and a master-slave PGA to improve efficiency.

However, parallel computing systems have limitations, including computational intensity during large population iterations, communication latency from independent evolution, potential compromises in convergence speed and solution quality with suboptimal subproblem division, lack of synchronization affecting convergence rates, and scalability challenges in complex, high-dimensional, and dynamic environments. Therefore, researchers are focusing on integrated computing systems to solve these problems.

2.1.4 Integrated computing systems for path planning and replanning

Integrated computing systems combine sequential and parallel processing, utilizing a hybrid approach that leverages the strengths of both paradigms (Garrett et al., 2021). This integration parallelizes specific algorithm components while maintaining a sequential structure for others, balancing simplicity and efficiency (Bertram et al., 2006). These systems aim to optimize computation for both speed and complexity (Gul and Rahiman, 2019), making them well-suited for real-time CALPP-RP systems in dynamic environments.

Panda et al. (2020) proposed integrating algorithms with environments for efficient path planning and replanning in autonomous underwater vehicles (AUVs). Woo et al. (2021) combined the P-RRT* algorithm with a line-of-sight path optimizer for rapid solutions. Bertram et al. (2006) integrated inverse kinematics and path planning for redundant manipulators. Zhu et al. (2022) integrated metaheuristic optimization with virtual prototyping, focusing on automatic construction planning and path visualization. However, achieving occlusion-free and collision-free paths is challenging with sequential computing, particularly

in complex and dynamic construction environments. Most robotic obstacle avoidance algorithms rely on expensive computational sampling or optimization-based planners with inflexible architecture. To address this issue, the authors emphasize refining occlusion and collision detection algorithms and incorporating integrated parallel computing for crane lift planning in dynamic construction environments. Despite acknowledging parallel computing limitations, the authors argued that an integrated computing system is optimal for achieving occlusion-free and collision-free crane lift path planning and replanning in dynamic environments. Therefore, this study aims to develop an integrated computing system for occlusion-free and collision-free crane lift path planning and replanning in dynamic environments, addressing computational time, communication, convergence speed, solution quality, convergence rate, and scalability issues.

2.2 Predicting degraded lifting capacity during lifting operations

2.2.1 Analytical approach for crane degradation

Analytical methods use tools, models, mathematical equations, and formulas to address complex problems. In crane assessment, these techniques are applied to evaluate and predict factors such as structural integrity, LC, and overall safety. This approach involves understanding failure mechanisms such as cracks, wear and tear, deflection, friction, and stress, identifying critical components, considering environmental conditions, and analyzing historical data. Relevant information regarding crane usage, environmental factors, and material properties is collected and organized for analysis. Monitoring indicators of early degradation throughout the crane's operational lifespan can help detect potential issues in advance.

Researchers have studied various components of the crane to investigate specific failure mechanisms. For instance, Vukelic et al. (2019) investigated the failure mechanism of crane

gear damage, while Guerra-Fuentes et al. (2020) studied the failure analysis of steel wire ropes in overhead crane systems. Das et al. (2018) reviewed the failure analysis of hooks, and Wu et al. (2021) developed models for fatigue damage accumulation in critical components exposed to moving loads. Experimental and numerical methods are employed to examine component behavior and identify the root causes of failure. Experimental techniques involve visual inspections, chemical composition analysis, mechanical properties, and hardness tests. Numerical methods, such as finite element analysis (FEA), are utilized to calculate stress/strain and load/deflection, providing insights into the structural performance of crane components.

However, it is essential to acknowledge that the analytical method has limitations. It requires a comprehensive understanding of each parameter used in the model and extensive experimental work to comprehend the failure mechanisms of individual components. Therefore, this method may be less effective for complex machines like cranes, as it relies on assumptions and suppositions regarding various parameters.

2.2.2 Data-driven method for crane degradation

A data-driven method (DDM) analyzes and interprets data to address problems and make informed decisions. It utilizes statistical, machine learning (ML), or artificial intelligence (AI) techniques to collect and scrutinize relevant data related to the degradation of tower cranes. The main goals of this approach are to monitor the operational state and performance of the cranes, predict potential issues or failures, and optimize maintenance and operational strategies. Various types of data are considered, including operational data (e.g., load, usage, and environmental conditions), sensor data (e.g., vibration and temperature), maintenance records, and historical failure data. A thorough analysis of this data is conducted to identify patterns and trends.

For example, Tran et al. (2012) conducted a study where they employed a support vector machine (SVM) to predict machine degradation and estimate the remaining useful life (RUL) of machines. They incorporated monitoring and operational data, along with degradation indices, to establish a failure threshold for decision-making. However, it is important to note that the effectiveness of the DDM approach relies on having sufficient data available, and without a solid understanding of the specific domain's deterioration mechanisms, its practical application may be less reliable. Additionally, the data in this particular study was sourced from only one or two sensors installed over the course of a year.

To overcome this issue, Roman et al. (2021) introduced a hybrid data-driven fuzzy active disturbance rejection control system for tower cranes. Similarly, Liu et al. (2021) utilized hybrid deep neural networks to assess machine performance based on fatigue data. The hybrid approach involves collecting data from multiple sensors, but it can result in complex and computationally intensive modeling due to the increased data volume. This is where the digital twin-driven (DTD) method comes into play, as it allows for real-time data acquisition, prediction of responses, and interaction between humans and machines.

2.2.3 Digital twin-driven method for crane degradation

A digital twin-driven (DTD) approach utilizes digital twin technology to monitor the operational condition, performance, and potential degradation of a complex physical system, such as tower cranes. A digital twin is a virtual replica of the physical object created through advanced data analytics, simulation, and modeling techniques. Researchers have applied DTD technology to address various machine performance degradation problems. For instance, Deebak and Al-Turjman (2021) used the DTD method for fault diagnosis in machining tool conditions. In contrast, Zhu et al. (2021) used it to control the machining process for thin-walled part manufacturing. Zhang et al. (2021) utilized DTD for surface roughness prediction, and Liu et al. (2021) employed it for traceability and dynamic control of machine processes.

Zhuang et al. (2021) utilized DTD to predict wear in the turning process of machines. Jiang et al. (2022) focused on the stability analysis of the tower crane hoisting process by examining stress conditions during various hoisting scenarios. Finally, Yang et al. (2022) used DTD for performance degradation in the transmission unit of CNC machine tools. However, their approach involved incorporating damage indices with wear data from a specific model, making it time-consuming, complex, and applicable only to that particular tool. Therefore, this study proposes a method for predicting degradation in aging tower cranes based on real-time data, offering a more efficient and versatile approach.

In summary, each method such as analytical, data-driven, and digital twin-driven offers distinct advantages and limitations when addressing the degradation of crane systems. Analytical methods provide insights into mechanical failure mechanisms but often depend on assumptions and intensive experimental validation. Data-driven methods leverage historical and sensor data to improve prediction accuracy, though their reliability can be compromised without sufficient domain knowledge or data quality. The digital twin-driven (DTD) approach has emerged as a transformative framework by enabling real-time monitoring, bidirectional interaction between physical and digital systems, and more responsive maintenance strategies. However, it is essential to recognize that digital twins do not replace fundamental failure prediction models; rather, they enhance them by integrating physical simulation, data analytics, and real-time IoT sensing. The effectiveness of a digital twin depends on the accurate modeling of failure mechanisms and the seamless fusion of stress/strain data, environmental variables, and material behavior. Thus, a holistic degradation prediction system should synergistically combine traditional analytical insights, data-driven analytics, and DT-enabled visualization and simulation to support reliable and robust crane health monitoring and decision-making.

2.3 Safety risk assessment using cascade learning

2.3.1 Construction safety management via computer vision

The construction industry is characterized by a high-risk environment where workers frequently encounter accidents, injuries, and fatalities (Shao et al., 2019). Factors contributing to these incidents include hazards associated with machinery proximity and unsafe operations (Duarte et al., 2021). Traditionally, safety inspections have been largely dependent on human oversight, which can be limited to both scope and on-site effectiveness (Stein, 2018). Consequently, there is an increasing need for automated and intelligent monitoring systems to enhance safety outcomes (Kolar et al., 2018). Computer vision is indispensable to achieve dynamic real-time perception and visualization of safety risks at construction sites. Computer vision, supported by deep learning, provides solutions by extracting and analyzing data from visual sources, such as videos, images, and CCTV footage (Liu et al., 2022; Moragane et al., 2024). These technologies are utilized for object detection, image classification, motion tracking, action recognition, and worker pose estimation (Nath et al., 2020; Son and Kim, 2021). They are crucial in performing detailed calculations and generating early warnings for detected risks. These functionalities are essential for understanding the behavior of machinery and workers, offering valuable insights into site activities and playing a vital role in detecting, assessing, evaluating, mitigating, and monitoring hazards, thereby improving safety protocols (Fang et al., 2020).

Computer vision also facilitates the automation of construction management tasks that typically require thorough inspection, such as safety management, quality control, progress monitoring, and productivity tracking (Paneru and Jeelani, 2021). For instance, Nath et al. (2020) implemented three deep learning models (i.e., VGG-16, ResNet-50, and Xception) based on the You-Only-Look-Once (YOLO) architecture to verify real-time compliance with personnel protection equipment (PPE) by detecting whether a worker is wearing a hard hat,

vest, or both. Similarly, Lin and Fang (2013) developed an automated tile installation quality assurance system using computer vision technologies. Reja et al. (2022) introduced a Computer-Vision-Based Construction Progress Monitoring (CV-CPM) system to enhance project monitoring. Kazemian et al. (2019) designed a vision-based real-time extrusion system for productivity tracking in robotic construction, addressing various levels of material variations. Additionally, computer vision can improve proximity-hazard detection; for example, Gu et al. (2022) developed an efficient human-machine collision warning system that uses computer vision and deep learning to measure the distance between machinery and workers in real-time based on specific thresholds. However, more research is necessary to develop effective methods for identifying unsafe equipment conditions on construction sites that could result in safety incidents.

2.3.2 Risk assessment in construction sites

The selection of assessment methodologies significantly influences the effectiveness of risk assessment outcomes in construction projects. The industry has evolved through various risk assessment approaches as researchers have addressed the sector's dynamic requirements. For instance, Jannadi and Almishari (2018) developed the Risk Assessment Model (RAM), which conceptualizes risk through three fundamental components: the probability of occurrence, the severity of hazards, and exposure levels at construction sites. Lee et al. (2012) established a weighted value system through a systematic literature review and questionnaire analysis to evaluate risk factors and their corresponding severity and frequency across various work categories. Arumugaprabu et al. (2022) introduced a quantitative approach to risk analysis by combining exposure frequency with legislative frequency using a comprehensive hazard source inventory. Furthermore, Sanni-Anibire et al. (2020) implemented a comparative survey methodology to evaluate accident risks and determine risk scores and weights through a weighted survey analysis. Lyu et al. (2020) implemented the Fuzzy Hierarchical Analysis

Process (FAHP) by developing fuzzy numbers and a judgment matrix based on specialized questionnaires. Gashaw and Jilcha (2022) proposed a comprehensive framework that integrates qualitative and quantitative elements through modified fuzzy integrated evaluation techniques to assess risk likelihood and impact. Similarly, Gunduz and Laitinen (2018) introduced an innovative approach to construction safety management using a control-level-based assessment method, which deviates from traditional probability metrics.

The integration of computer vision technology has enabled real-time safety alert systems through proximity warnings, providing proactive notifications to both ground workers and equipment operators. Advancements in active warning technologies have enhanced the ability of operators to implement timely hazard avoidance strategies. Applications of visual technology have strengthened workers' ability to recognize risks associated with various incidents, including explosions, falls from height, and vehicle-human interactions (Wu et al., 2022). A notable advancement by Hou et al. (2023) involved the implementation of computer vision to identify hazard sources and nearby objects. Their risk calculations were based on the Swiss Cheese model and Heinrich's causal chain methodology, incorporating questionnaire-based impact assessments for comprehensive risk evaluation.

Construction site safety was significantly enhanced by analyzing workers' visual inspection patterns, which demonstrated a strong correlation with hazard recognition capabilities. Integrating eye-tracking technology with computer vision systems facilitates automated analysis of machine movements, providing operators with personalized feedback to enhance risk management. Safety risk monitoring is optimized through early warning systems that continuously visualize potential hazards on management platforms, enabling site supervisors to control hazards in real-time (Liang and Liu, 2022). The implementation of computer vision and deep learning technologies has advanced safety measures and regulatory compliance and substantially improved the industry's capability for object detection and tracking.

2.3.3 Object detection and tracking

Object detection and tracking play a vital role in the construction industry by enhancing safety monitoring by identifying and tracking machinery and worker movements. This proactive approach helps to prevent accidents and improve overall site safety (Kulinan et al., 2024). Various strategies have been proposed to tackle these challenges. For instance, Xiao and Kang (2020) introduced the Alberta Construction Image Dataset (ACID), which includes 10,000 images of 10 construction machines. This dataset significantly reduces the manual effort required for annotations and is a benchmark for validating different detection algorithms such as YOLO-v3, Inception-SSD, R-FCN-ResNet101, and Faster-RCNN-ResNet101. Similarly, Kim et al. (2018) employed a region-based, fully convolutional network coupled with transfer learning to detect construction equipment like Excavators, Dump trucks, Loaders, Road rollers, and Concrete mixers. Their approach aids in improving site management and planning by accurately recognizing construction equipment using both real job site images and benchmark data.

Additionally, Lee et al. (2022) utilized a deep learning-based object detection algorithm, specifically the Faster R-CNN model, to detect construction equipment. They used synthetic datasets generated from the Unity 3D game engine, incorporating virtual objects into real-site images to reduce the time and effort needed for training data collection and annotation. In more recent research, Shin et al. (2024) applied the YOLOv5 model to automate the detection and classification of nine different types of construction equipment, achieving a mean average precision (mAP) of 90.26%, which outperformed results from the Microsoft Common Objects in Context (MS COCO) dataset.

Further, lightweight models have been developed for real-time object detection. For example, Chen et al. (2023) introduced a face-assisted model (WHU-YOLO) that integrates a ghost module and a Bi-directional Feature Pyramid Network (Bi-FPN) for object detection. Other

models, such as YOLO-NL and Augmented Weighted Bidirectional Feature Pyramid Network (AWBiFPN), have been applied to marine object detection, utilizing global strategies for dynamic label distribution while enhancing the Cross-Stage Partial Network (CSPNet) and Path Aggregation Network (PANet) with gradient strategies and attention mechanisms (Gao et al., 2024; Zhou, 2024).

These advancements underscore the ongoing progress in developing precise, efficient, and context-specific object detection methods suitable for complex construction environments. However, further research is required to fully implement computer vision for safety management, particularly in offsite construction, where the risks posed by moving objects such as mobile and tower cranes remain significant (Alsakka et al., 2023). Continued innovation in object detection and tracking is crucial for improving construction safety monitoring and extending its applications to areas such as activity recognition and analysis.

2.3.4 Activity recognition, classification, and assessment analysis

In recent years, there has been growing interest in leveraging computer vision and deep learning techniques to enhance the recognition, classification, and analysis of activities in construction site environments (Fang et al., 2020c; Khallaf and Khallaf, 2021). Various learning approaches, including supervised, semi-supervised, unsupervised, self-supervised, and reinforcement learning, have been explored in this context. For instance, Pandey et al. (2024) introduced a method that combines a frequency-based blur-variance local phase quantization feature extractor with a multiclass SVM classifier to recognize human activities. Similarly, Khowaja et al. (2024) presented the Representation Fusion using Self-supervised Learning for Activity Recognition (ReFuSeAct) framework, which effectively classifies human activities using modality-specific encoders, attention mechanisms, and decision-level fusion strategies with minimal annotated wearable sensor data. Another approach by Martin et al. (2023) proposed a three-stage architecture integrating a VGG16 feature extractor, a Transformer Encoder-Long

Short-Term Memory (LSTM) module, and post-processing components to recognize and localize excavator actions. Liu et al. (2022) focused on detecting and localizing unsafe worker behaviors in digital images using a visual grounding technique. In addition, Fang et al. (2018) applied computer vision to monitor the load sway of cranes on offshore platforms.

Recent advancements in neural networks, particularly in cascade learning and the integration of transformers, have led to significant improvement in activity recognition, classification, and analysis within construction sites. Chian et al. (2022) proposed a dynamic key point identification method for estimating crane load and fall zones. Zhang and Ge (2022) developed a dynamic collision prewarning mechanism for tower cranes, utilizing FairMOT-based tracking algorithms in conjunction with transformers. Furthermore, Jiang and Ding (2024) employed transfer learning methods to identify unsafe hoisting behaviors in tower cranes, while Kim et al. (2024) used the cascade learning method to detect unsafe operations of construction machinery, including Backhoes, Bulldozers, Dump trucks, Excavators, and Loaders.

2.3.5 Limitations and technical challenges in current studies

Recent advancements in computer vision (CV) and deep learning have significantly enhanced safety management on construction sites. These technologies have shown considerable potential in improving the ability to monitor, detect, and mitigate hazards. However, despite these advancements, several challenges persist, particularly with regard to data scarcity and the accuracy of CV algorithms under varying site conditions.

One of the primary limitations is the lack of comprehensive, well-annotated datasets specific to the construction industry. The process of data collection for large construction equipment, such as cranes, presents unique challenges due to the limited availability of online resources and the impracticality of simulating unsafe operations. Moreover, capturing multiple cranes within a single frame, along with dealing with issues of data quality, further complicates

research efforts. Current datasets, including MOCS (Xuehui et al., 2021), MCER (Tajeen and Zhu, 2014), SODA (Duan et al., 2022), ACID (Xiao and Kang, 2020), and synthetic image datasets (Lee et al., 2022), as well as those designed for human activity recognition, like COCO (Fleet et al., 2014), UCF101 (Pandey et al., 2024), HMDB51 (Pandey et al., 2024), and SHWD (Wang et al., 2021), exhibited limitations in terms of scope, annotation depth, and generalizability. These limitations hinder the development of robust CV models that can generalize well across various construction environments.

In addition to data scarcity, the application of CV and deep learning in construction safety management is affected by site-specific and environmental factors. The detection of small objects, such as tools or workers, is often complicated by the size and positioning of objects within the camera's view, as well as the limited resolution and coverage of the cameras. Furthermore, variations in lighting conditions—common in outdoor and dynamic construction environments—can impact the accuracy of CV models. The task of capturing large machinery, such as cranes, is further challenged by the need to operate within the constraints of camera angles and the limited field of view, which may not adequately capture the full scope of activity on site. Additionally, the requirement for real-time processing of high-resolution images necessitates the optimization of models to adapt to the diverse and dynamic conditions inherent in construction sites.

Another key challenge lies in the precision of pose estimation models, which may suffer from inaccuracies when trained on non-specific datasets. Automated detection systems can struggle with variations in object scales and environmental conditions, and lightweight models, while optimized for real-time performance, often lack the robustness and generalization capabilities necessary for accurate detection in complex construction environments. Furthermore, multi-object tracking, particularly when multiple cranes or machinery are present in the same frame,

presents a significant challenge. Models require integrated frameworks to enhance accuracy, especially in detecting irregular activities or hazardous behaviors. Real-time deployment and scalability remain obstacles for automated hazard identification systems, which often face difficulties in handling complex, dynamic scenarios.

To address these challenges, this study proposes an innovative model that integrates advanced transformer-based models in a cascaded approach. This cascaded transformer-based model leverages a comprehensive dataset, which includes hazard detection involving both mobile and tower cranes sourced from various YouTube videos. The model enhances low-resolution video footage using Super-Resolution Generative Adversarial Networks (SRGAN) and applies the Real-Time DEtection Transformer with Locality (RT-DETR-L) model for real-time object detection. By combining SRGAN with RT-DETR-L, the proposed cascade learning method not only improves detection accuracy but also reduces the occurrence of false positives. Furthermore, a hierarchical classification system has been incorporated to enable effective real-time monitoring and response, facilitating the detection and classification of objects and actions for proactive safety management on construction sites.

While the proposed model significantly improves detection accuracy, it is essential to recognize that CV algorithms' performance can vary significantly depending on the scenario and environmental conditions. Factors such as lighting variations, camera angles, and object scale can impact the accuracy of CV models, leading to inconsistent results across different construction sites. For example, detecting small objects in cluttered environments remains a significant challenge due to limited camera coverage and resolution. Similarly, detecting moving machinery, such as cranes, is complicated by the dynamic nature of construction sites, where multiple machines and workers are often present simultaneously.

Furthermore, data scarcity remains a critical issue in the construction industry. The availability of large, well-annotated datasets, particularly for training CV models on construction machinery like cranes, is limited. The shortage of such datasets is exacerbated by the difficulty of capturing unsafe operations or hazardous scenarios in real-life datasets, which are essential for training robust predictive models. Existing datasets, such as MOCS, MCER, SODA, and ACID, exhibit significant limitations in coverage, annotation depth, and generalizability to various construction sites. Therefore, the development of more comprehensive and diverse datasets tailored specifically to construction safety, along with the optimization of models to handle real-time data from diverse site conditions, remains an important area for future research.

2.4 Chapter Summary

Crane operations in dynamic construction environments face significant safety challenges due to moving obstacles, structural fatigue, and complex site layouts. Several optimization algorithms, including deterministic, probabilistic, and evolutionary methods, have been explored for path planning and replanning. Hybrid approaches, such as PSO-SA and GA-SA, have also shown promise in improving efficiency and adaptability. However, sequential computing systems, while effective for static environments, struggle with scalability in dynamic settings due to their linear processing structure. Parallel computing systems, which offer improved speed through concurrency, encounter difficulties in synchronization and suboptimal subproblem division. Integrated systems that combine both sequential and parallel paradigms show potential in balancing speed and complexity but remain underexplored for achieving occlusion-free and collision-free path planning in dynamic environments. Key gaps in the field include the absence of scalable frameworks for real-time replanning in cluttered, evolving environments. Additionally, there is a critical need for advanced detection algorithms capable of addressing dynamic occlusions that emerge during crane operations. While sensor-

based tracking systems such as GPS, UWB, and IMU, as well as 4D BIM integration, enable collision avoidance and progress monitoring, these technologies are often limited by the lack of cohesive, real-time decision-support systems. Furthermore, constraint-based modeling for spatial feasibility assessment is hindered by interoperability challenges between sensing, modeling, and analytics platforms, restricting their full potential in dynamic environments.

In relation to aging cranes, structural degradation driven by factors such as fatigue, corrosion, and cyclic loading presents significant safety risks. Analytical methods, including Finite Element Analysis (FEA) and experimental testing, provide mechanistic insights but are based on assumptions and require extensive validation. Data-driven approaches, such as Support Vector Machines (SVM) and hybrid neural networks, utilize sensor and historical data to predict degradation but rely on long-term datasets and domain-specific knowledge to ensure reliable predictions. Digital twin (DT) technology has emerged as a tool that integrates real-time sensor data with virtual models, enabling the simulation of degradation processes. However, current DT frameworks for cranes are narrowly focused—typically on specific components or tools—and lack validation for long-term (0–70 years) lifting capacity forecasting. There remains a critical gap in integrating analytical failure models with real-time DT-enabled degradation monitoring to create practical, scalable solutions for aging crane safety.

In the realm of safety management, computer vision (CV) and deep learning techniques have advanced the detection of unsafe crane operations. These techniques facilitate object detection, activity recognition, and proximity warnings. However, challenges persist regarding dataset quality (e.g., MOCS, ACID), model generalizability, and real-time performance under dynamic site conditions. Lightweight models, such as YOLO variants, are effective in balancing speed and accuracy but struggle with small-object detection, occlusions, and tracking multiple cranes

simultaneously. More sophisticated models, such as hierarchical classification and transformer-based networks (e.g., RT-DETR-L), improve detection precision but require further optimization for construction-specific hazards. The main unresolved challenges in this area include: (1) Data Scarcity: The lack of annotated datasets for unsafe crane operations and hazardous scenarios, (2) Environmental Variability: The inconsistency of model performance under varying lighting conditions, camera angles, and cluttered backgrounds, and (3) Real-Time Scalability: The computational bottlenecks involved in processing high-resolution video footage for proactive risk mitigation.

CHAPTER 03

INTEGRATED COMPUTING SYSTEM FOR OCCLUSION-FREE AND COLLISION-FREE PATH PLANNING AND REPLANNING FOR ROBOTIZED CRANES IN DYNAMIC ENVIRONMENT

3.1 Introduction

In modern construction, prefabricated volumetric modules are frequently produced off-site and transported for on-site installation (Kayhani et al., 2021; Taghaddos et al., 2018; Zhang and Pan, 2021; Zhou et al., 2023). Traditionally, certified operators use tower cranes to lift these modules from loading areas to their final positions (Al-Hussein et al., 2006; Ji and Leite, 2018; Lin et al., 2014). However, this blind lifting operation relies heavily on human judgment, making it inefficient and risky. It can potentially cause collisions with the surrounding obstacles, leading to injuries, fatalities, and financial losses (Tam and Fung, 2011; Zhou et al., 2018a, 2018). In 2022, Hong Kong reported 3,065 construction accidents, including 47 crane-related accidents, an increase from 39 in 2021. Of these crane accidents, 38% occurred during operation, 31% during assembling or dismantling, and 23% due to other factors (Gu et al., 2021; Sadeghi et al., 2021). Collision with surrounding obstacles remains the leading cause of crane accidents during operation (Ajmal et al., 2016; Ali et al., 2005). The critical importance of crane safety highlights the need for computer-aided lift path planning and replanning (CALPP-RP) to enhance current safety measures in crane operations. CALPP-RP requires rapid and accurate computation of real-time collision-free lift paths.

CALPP-RP systems employ advanced algorithms, intelligent decision-making, computer graphics, and simulations for real-time adaptation in construction projects (Chang et al., 2012; Hu et al., 2021). They provide collision-free optimal lifting paths for cranes, thereby increasing safety and productivity (Ajeil et al., 2020). Current CALPP-RP systems often treat planning and replanning as separate problems, resulting in suboptimal solutions. In dynamic

construction scenarios, tower cranes operate in constantly changing environments that require continuous obstacle avoidance (Lei et al., 2013; Soltani et al., 2002). Planning algorithms must incorporate replanning steps to adjust initial lifting paths according to environmental changes (Li et al., 2020; Zhang and Hammad, 2012). Real-time replanning algorithms are crucial for swiftly modifying affected path segments based on new information (Dakulovi and Petrovi, 2011; Zhang and Hammad, 2012). Therefore, a real-time, automated, and integrated system for collision-free crane lift path planning and replanning in dynamic environments is essential.

Various methods have been explored in CALPP-RP systems to achieve automated and real-time lift path planning and replanning (Hu et al., 2021). These methods can be broadly categorized into deterministic, probabilistic, and evolutionary algorithms. Deterministic algorithms, such as Dijkstra and Dynamic A*(D*), find paths between start and end configurations by following predefined rules and exploring the solution space, often using heuristics to determine the optimal paths (Koenig et al., 2004; Soltani et al., 2002). These algorithms offer efficient adjustments to previously planned paths in response to environmental changes (Ferguson and Stentz, 2016; Stentz, 1994). However, they have limitations, including replacing all infeasible configurations within affected path segments without considering the potential advantages of alternative paths. Furthermore, they can be computationally intensive and lack adaptability, often necessitating repeated recalculations of optimal paths. Researchers have proposed heuristic techniques to enhance the efficiency of deterministic algorithms. For instance, Stentz (1995) introduced an extension to D* that employs a heuristic focusing function based on the position of the target load, resulting in reduced runtime. Another modification, the Two-Way D* (TWD) algorithm, utilizes a 2D occupancy grid map (Dakulovi and Petrovi, 2011). Nevertheless, as the dimensions of the configuration space (c-space) increase, deterministic algorithms become less feasible due to their computational demands.

Probabilistic algorithms, such as Simulated Annealing (SA), Stochastic Gradient Descent (SGD), and Rapidly Exploring Random Trees (RRTs), find paths between start and end configurations by exploring the solution space using random processing (Hu and Fang, 2016; Zhang and Hammad, 2012). These algorithms are versatile and valuable in various motion-planning scenarios. However, navigating the high-dimensional c-space of cranes in complex environments using random sampling can be challenging. It is essential to note that probabilistic algorithms do not guarantee finding the globally optimal path (Zhou et al., 2021). To address these challenges, researchers have proposed modifications. For instance, Bruce and Veloso (2014) introduced the Extended RRT (ERRT) to balance planning and executing lifting paths in real-time. Ferguson et al. (2014) implemented the Dynamic RRT (DRRT) approach, which efficiently prunes only the invalid segments of the initially planned path rather than discarding it or starting anew. AlBahnassi and Hammad (2012) contributed to the field with a framework for near real-time crane motion planning, employing a modified DRRT strategy tailored for efficient planning in dynamic environments. Miao and Tian (2013) extended the capabilities of standard SA by introducing path operator deletion, switching, and incorporating heuristic-based initial path selection. Nonetheless, it is worth noting that their algorithm focuses on adjustments within the RRT that are affected by biased probabilities. A challenge arises when determining the optimal level of bias during the planning phase, as an excessively high biasing probability can lead to inefficiencies in DRRT operations (Anwar et al., 2024, 2023; Hussain et al., 2024, 2023; Zheng et al., 2023).

Evolutionary algorithms, such as Genetic Algorithm (GA), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO), find paths between start and end configuration by mimicking the process of natural evolution (Aslan et al., 2023; Tuncer and Yildirim, 2012). These algorithms operate on a population of potential solutions, refining them through processes analogous to selection, crossovers, and mutation. For instance, Mohajer et al. (2013)

developed an online random PSO for mobile robots navigating dynamic environments. However, there has been no significant improvement in the computational complexity of path planning and replanning. Researchers have attempted to use hybrid algorithms to achieve real-time path planning and replanning without additional costs. For instance, Yang Wang (2007) proposed a hybrid GA approach focused on obstacle vertices in the search space and incorporated robot speed selection into the genetic framework, resulting in a broader range of potential solutions. However, this hybrid algorithm fails to achieve real-time path planning and replanning.

Researchers have recently explored various computing techniques to implement and execute optimization algorithms sequentially and in parallel, aiming to achieve the planning and replanning in real-time. For instance, Zhu et al. (2022) presented a hybrid metaheuristic algorithm that combined PSO with SA sequentially to facilitate safe and efficient crane lift path planning. Cai et al. (2018) used a prioritized multi-objective parallel genetic algorithm (PGA) for automatic path planning for dual cranes in dynamic environments. Similarly, Dutta et al. (2020) utilized a GPU-based parallelization method for path replanning, leveraging a multilevel depth map (MDM) and a master-slave parallel GA to enhance efficiency.

Despite these valuable contributions, several research gaps need to be addressed to improve the computational efficiency of robotized tower crane lift path planning and replanning in dynamic environments: (1) delays arise from elevated computational intensity due to communication overhead and load imbalances; (2) convergence rates are impacted by synchronization issues, load balancing, and algorithmic dependencies; (3) the success rate is hindered by a constrained exploration of the solution space; (4) instability is introduced through suboptimal subproblem division; (5) the system grapples with trade-offs between path smoothness and computational efficiency; and (6) challenges in communication latency and scalability emerge in intricate, high-dimensional, and dynamic environments. To address these

gaps, this study introduces an integrated approach for occlusion-free and collision-free crane lift path planning and replanning in dynamic construction environments.

3.2 Methodology

3.2.1 Overview of research methodology

This section outlines our research methodology, employing qualitative and quantitative methods to develop and evaluate an integrated computing system for occlusion-free and collision-free crane lift path planning and replanning in dynamic construction environments. As illustrated in Figure 3.1, the methodology comprises four main steps: (1) environment and simulation platform; (2) optimization platform; (3) evaluation platform; and (4) development of the decision support system (DSS) and Path Re-planner (PRP) system. We explain each component's necessity, subcomponents, functionality, inputs, and outputs within the integrated system. Then, we describe how these components are unified to manage planning and replanning tasks in complex construction scenarios.

The environment and simulation platform describes the workspace development for the integrated computing system. We created 3D models, designed the environment, and implemented physics simulations for crane lift using Unity. Unity was chosen for this study due to its robust physics engine, real-time simulation capabilities, support for dynamic environments, and seamless integration with custom algorithms via C# scripting. These features enabled accurate modeling of crane operations, including interactions with both static and dynamic obstacles commonly found on construction sites. While real-world data was not directly used, simulation parameters were derived from actual crane specifications and safety protocols to ensure realism. The virtual environment also included dynamic entities such as moving equipment and personnel to reflect practical site conditions. Although real-world validation remains a future goal, Unity provided a flexible and effective platform for

developing and evaluating the proposed system under complex, near-real scenarios. We appreciate your comment, which helped us clarify this justification. We transformed the workspace into a configuration space (c-space) using robotics principles (Dutta et al., 2020; Li et al., 2023; Li et al., 2020; Luan and Thinh, 2023; Lynch and Park, 2016; Nearchou and Aspragathos, 1997; Razali et al., 2023). This step provides inputs to the optimization platform and receives optimal solutions for DSS and PRP systems.

The optimization platform describes the mathematical modeling of a hybrid algorithm combining A* with a Genetic Algorithm (GA). A* searches the c-space to generate solutions, which GA then optimizes. TOPSIS selects optimal and suboptimal paths, while raycasting detects occlusions and collisions. C# was preferred over MATLAB and Python for faster execution, better memory management, multithreading capabilities, and integration with optimized libraries (Li et al., 2020; Roberge et al., 2013). Both platforms require validation to reflect real-world scenarios accurately.

The evaluation platform describes the testing and validation of the simulation and algorithm platform. A pilot project in Unity served as a case study to assess the effectiveness and efficiency of both platforms. The case study involved four symmetrical 4-story building blocks with over 100 prefabricated modules, representing typical challenges in building projects. The evaluation focused on the 4D crane lift simulation and optimization algorithms platforms. The simulation platform was assessed through self-examinations and semi-structured interviews with industry practitioners, addressing function testing, realism, visualization, animation, performance, operability, scalability, communication, interoperability, and data integration. The algorithms' evaluation aimed to assess the hybrid algorithm's effectiveness and efficiency across four crane lift scenarios, including planning and replanning. Performance analysis considered computational time, convergence speed, success rate, solution stability, and path smoothness.

Finally, a DSS was developed to provide real-time guidance and support decision-making for operators and project managers, enhancing crane lift operation efficiency and safety. It included a collision warning system alerting users to potential hazards based on proximity information and predefined thresholds. The PRP functioned as a dynamic system, automatically recalculating optimal crane lift paths in response to environmental changes or collision risks. The DSS and PRP were integrated into a user-friendly graphical interface for effective and intuitive control.

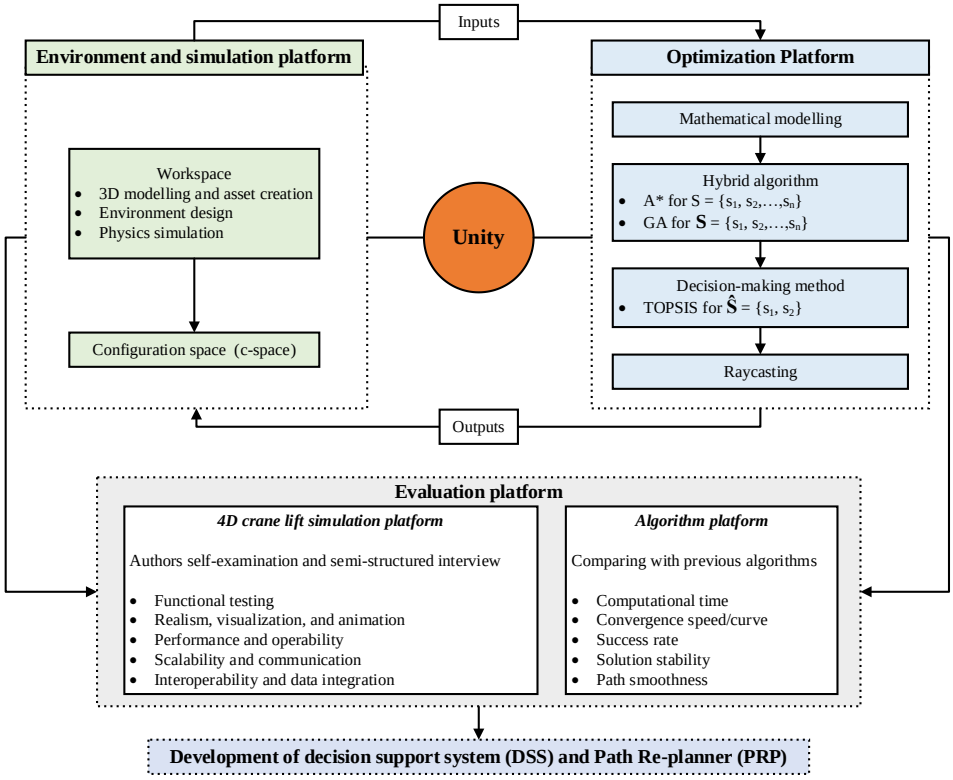


Figure 3. 1: Overview of research methodology.

3.2.2 Problem formulation

An integrated computing system for occlusion-free and collision-free lift path planning and replanning was developed for a robotized tower crane based on the assumptions of Dutta et al. (2020) and Zhu et al. (2022). While considering the characteristics of real-life construction projects. The system is based on the following assumptions:

1) A robotic tower crane with five degrees of freedom (DOFs) was utilized.

2) Static obstacles for discrete collisions and dynamic obstacles for continuous collisions were addressed.

3) It was ensured that no obstacles were present above the tower crane.

4) Crane operations were decoupled, allowing only one link to move at a time, which created an ordered lifting path.

5) The lifted module remained parallel to the ground, preventing rotation around the x- and y-axes.

6) The hoisted module underwent self-rotation only at the start and end positions.

3.2.3 Degree of freedom and configuration space

This study considers a specialized T-type tower crane with five degrees of freedom (DOFs), as depicted in Figure 3.2. Standard T-type cranes have three DOFs: sling hoisting, trolley movement, and jib swinging, which are based on revolute and prismatic joints. This study introduces two additional DOFs: self-rotation of hoisting modules and crane movement along its track. These extra DOFs expand the applicability to other crane types and facilitate scalability assessment in high-dimensional environments.

Assumption IV in Section 3.2.2 states that the path of a robotized tower crane comprises a sequence of elemental movements along defined axes. Each movement occurs independently along one axis within the crane's actuator local coordinate system. For better understanding, the workspace of the construction project is transformed into a configuration space (c-space). In c-space, crane positions for path planning and replanning are represented as configurations. The complete set of configurations in the solution space is denoted as X. Configurations in the *ith* path can be described by Eq. (3-1):

$$X^i = (X_{i1}, X_{i2}, X_{i3}, X_{i4}, X_{i5}) \tag{3-1}$$

where X_{ij} is a vector with N elements, $i \in [1, M]$, $j \in [1, 5]$, M is the population size or number of paths in the entire c-space, and N is the number of configurations in each path. Any configuration can be expressed using Eq. (3-2):

$$C = (\theta_s, \theta_{lr}, d_{tm}, d_h, d_{cm}) \tag{3-2}$$

Here, θ_s represents jib swinging, θ_{lr} is load rotation, d_{tm} is trolley movement, d_h is hoisting action, and d_{cm} is crane movement.

These configurations are linked by abstract connections called "edges," denoted as E. The set of edges in the *ith* path is labeled as E^i. Each edge represents the tower crane's movement between configurations. The lifting path is expressed as a sling, $s \subset S$, where S is the solution space of all paths, given by Eq. (3-3):

$$S = X \cup E \tag{3-3}$$

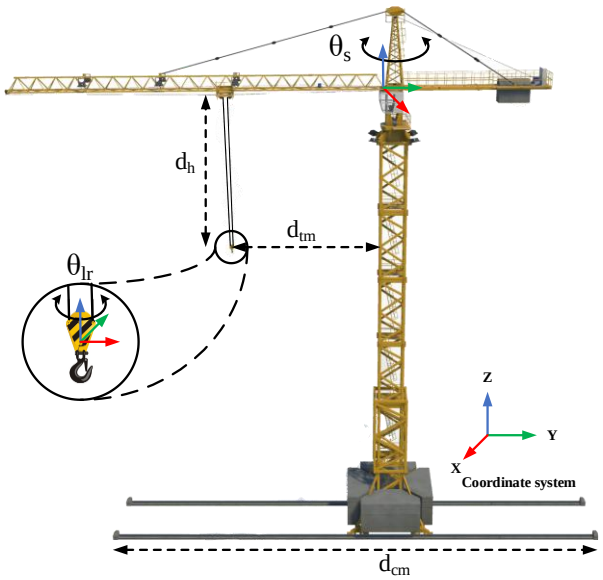
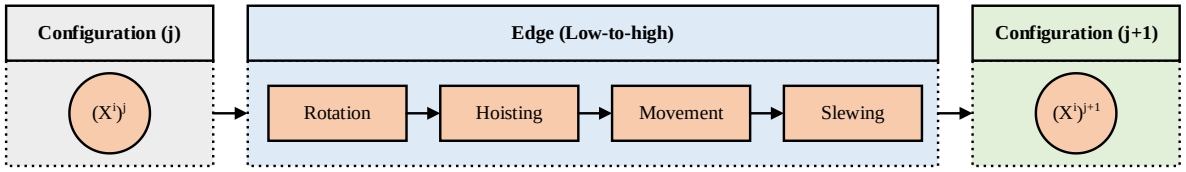


Figure 3. 2: Degrees of freedom of the chosen tower crane.

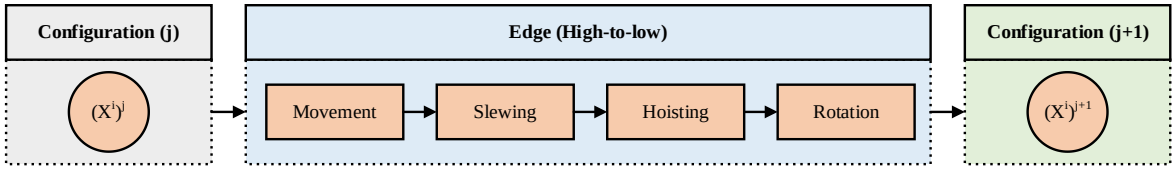
3.2.4 Sling formation by operational scenarios

Section 3.2.3 explored the interconnection of adjacent configurations via abstract edges representing crane movement. Within the overall path formation, three configurations are paramount: the start, the end, and those through which the lifting module traverses. Path formation delineates the operational sequence for crane navigation through the abstract edge between adjacent configurations, distinguishing low-to-high and high-to-low scenarios. This study focuses exclusively on low-to-high movement. In the low-to-high scenario, crane operations follow a specific sequence: Rotation, Hoisting, Movement, and Slewing. Conversely, the high-to-low scenario inverts this sequence. Figure 3.3 illustrates these operation sequences for each edge in the *ith* path, adapted from Zhu et al. (2022).

Let $(X^i)^j$ and $(X^i)^N$ denote the start and end configurations of the *ith* path, respectively. In practical crane operations, operators typically adjust the position of the lifting module by rotating it to the desired orientation before hoisting. Therefore, for $(E^i)^1$, the first edge introduces random variations in "Rotation" and "Hoisting" while maintaining other DOFs constant. As lifting progresses, the hook descends with the module. Subsequently, riggers rotate the module again for precise positioning at the target location $(X^i)^N$, typically from a fixed height (h). To accommodate this sequence, we introduce a pseudo-end configuration. The end configuration $(X^i)^{end}$ of the planned path is consistently set at a higher than the actual final position, ensuring stochastic changes only in "Rotation" and "Hoisting" for $(E^i)^{N-1}$. Maintaining these variables as separate edges guarantees their strict sequential order, while parameters for intermediary edges undergo random variations. This operational strategy for the tower crane employs a hybrid algorithm combining A* and GA.



(a) Low-to-high



(b) High-to-low

Figure 3. 3: Operation sequence for each edge in the ith crane lift path.

3.2.5 Hybrid algorithm

This study used a hybrid algorithm combining A* and genetic algorithms in an integrated system to address occlusion-free and collision-free crane lift path planning and replanning, as depicted in Figure 3.4. A* and genetic algorithms are distinct problem-solving approaches; their integration leverages the strengths of each method in addressing complex problems (Li et al., 2020).

3.2.5.1 A Algorithm*

The A* algorithm is a pathfinding and graph traversal method that finds the shortest path between start and end configurations in c-space. It begins by determining the start $\left(X^i\right)^1$ and end configuration $\left(X^i\right)^{end}$. A* then searches for solutions (S) using heuristic-guided search. A* is commonly used for pathfinding in c-space, where configurations represent position and orientation, and edges (E) represent transitions between adjacent configurations (Dakulovi and Petrovi, 2011; Soltani et al., 2002; Stentz, 1995). Multiple paths (S) are generated, denoted as $S = \{s_1, s_2, \ldots, s_n\}$.

To enhance the heuristic information and accelerate the convergence of GA, A*'s evaluation function (f) and bending suppression are introduced (Li et al., 2020). The f combines two

components: (1) the cost (g) to reach the $\left(X^i\right)^{end}$ from $\left(X^i\right)^1$, and (2) an estimated cost (h) from $\left(X^i\right)^1$ to $\left(X^i\right)^{end}$. Thus, $f = g + h$. In bending suppression, the A* selects the path (s) with minimum turns by choosing the configuration with the lowest f-score. These operators improve the quality of solutions (S).

However, A* alone has some limitations. It may not explore the entire search space, potentially missing better paths in complex scenarios. The accuracy of the heuristic estimate may not guarantee an optimal path, and A* might become trapped in local optima. Combining A* with the GA addresses these limitations. Therefore, set $S = \{s_1, s_2, \ldots, s_n\}$ serves as the initial population for the GA algorithm.

3.2.5.2 Genetic Algorithm

A GA, inspired by natural selection and genetics, finds approximate solutions to optimization problems by mimicking natural evolution. The fitness function $(\boldsymbol{f})$, based on velocity (v) and time (t), evaluates path planning consistency by assigning fitness values to each s. Selection, crossover, mutations, insertion, and deletion operators generate new smooth paths $(\boldsymbol{S})$ from those generated A*. The path with the highest $\boldsymbol{f}$ value is considered optimal.

Each s represents the solution space $\boldsymbol{S}$ of all paths in c-space. $\boldsymbol{S}$ has $N\ x\ 5$ dimensions, where N is the number of configurations in a crane lift path. Each solution s contain position $\left(X^i\right)$, velocity $\left(V^i\right)$, and traversal time (t) of the *ith* configurations. The position of the *ith* configuration is $X^i = (X_{i1}, X_{i2}, X_{i3}, X_{i4}, X_{i5})$. Similarly, the velocity for the *ith* configuration is denoted as $V^i = (V_{i1}, V_{i2}, V_{i3}, V_{i4}, V_{i5})$, where $i \in [1, M], j \in [1, 5]$.

The variables have specific ranges: (1) position: $X \in [X_{min}, X_{max}]$; (2) velocity: $V \in [V_{min}, V_{max}]$. X_{min} and X_{max} represent minimum and maximum values among tower crane

DOFs, while V_{min} and V_{max} correspond to minimum and maximum velocities allowed among DOFs, as shown in Figure 3.2.

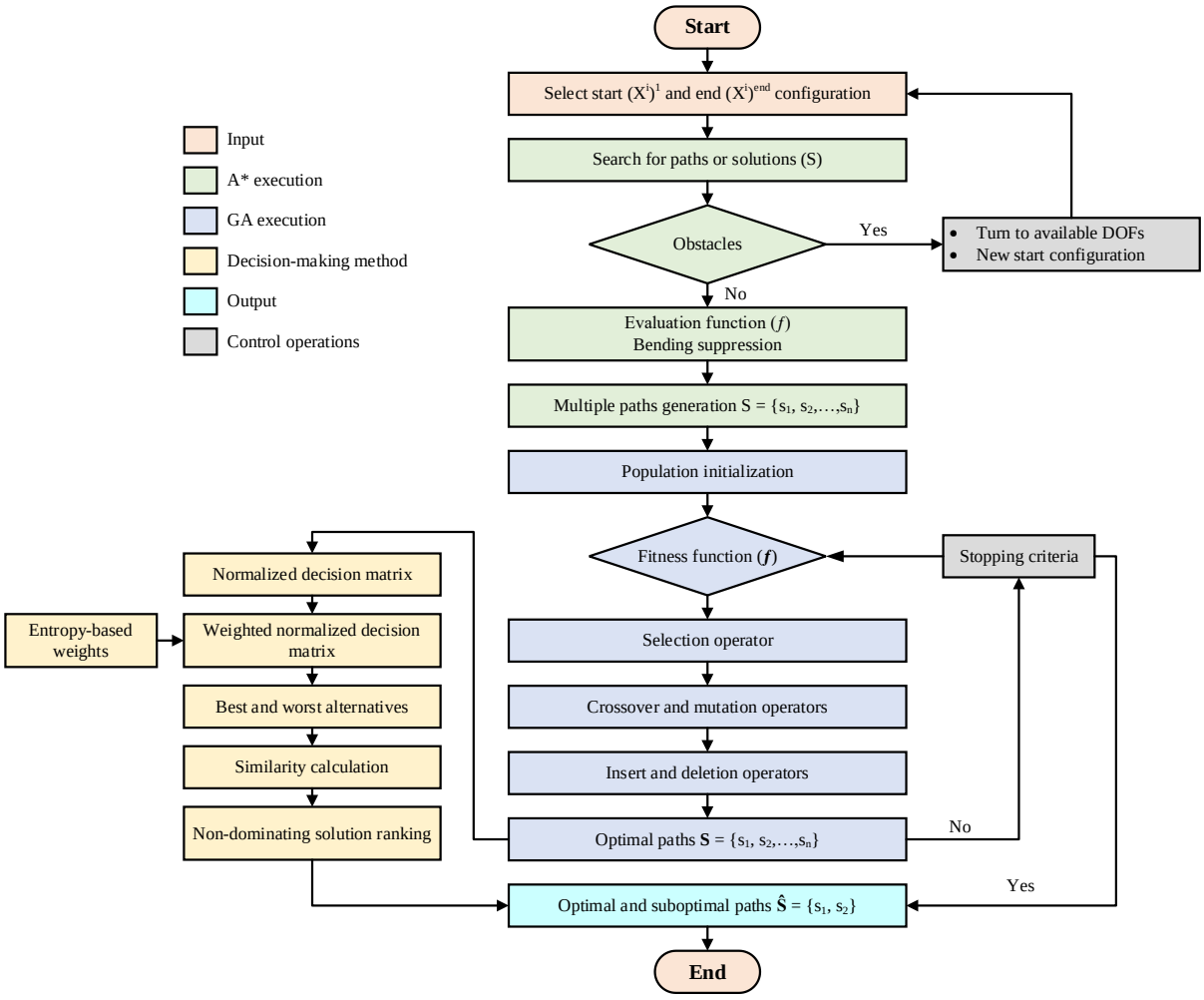


Figure 3. 4: Workflow for a hybrid algorithm.

3.2.5.3 Population initialization

The hybrid algorithm for path planning and replanning employs a two-step process to optimize paths in complex and dynamic environments. initially, A* searches the entire c-space, generating an initial population of paths $S = \{s_1, s_2, \ldots, s_n\}$. Subsequently, GA optimizes path selection. Using A* to generate the initial paths narrows the search space for GA, enhancing efficiency. GA iteratively selects the optimal path using a fitness function $(\boldsymbol{f})$ that evaluates efficiency, safety, and other relevant factors. The optimal paths are expressed as $\boldsymbol{S} = \{s_1, s_2, \ldots, s_n\}$. Initializing GA with A*-generated paths combines heuristic search and

optimization capabilities. This method streamlines path planning and replanning for tower crane operations, facilitating optimal solution identification in complex c-space scenarios.

3.2.5.4 Encoding of chromosomes

In path planning and replanning, each s represents a chromosome comprising three essential configurations: start, end, and intermediate. These configurations are termed "genes" within the chromosome. Two coding methods, Binary-Coded String (BCS) and Decimal-Coded String (DCS), create chromosomes (Ali et al., 2005; Tuncer and Yildirim, 2012; Zhang and Pan, 2020). BCS uses binary digits [0, 1] for discrete decision variables, whereas DCS uses decimal numbers for continuous variables. The problem involves abstract edges with constant parameters; DCS is preferable, offering greater flexibility and efficiency. A valid chromosome represents a path from start to end configuration. The GA-generated path set ($\boldsymbol{S}$) replaced the A*-generated set (S). These new paths result from the GA optimization of the initial solutions of A* to enhance stability and quality.

3.2.5.5 Fitness function

Path planning and replanning aim to find the optimal path between start and end configurations. A fitness function quantifies solution suitability. Optimal paths can be defined by criteria such as shortest distance, least time, lowest cost, or minimum energy consumption (Precup et al., 2022; Tubaileh, 2016). Typically, path planning problems focus on finding the shortest path (Tuncer and Yildirim, 2012). In this study, the fitness function ($\boldsymbol{f}$) of GA for a chromosome is represented in Eq. (3-4). The fitness value $\boldsymbol{f}_{(s)}$ is computed for a solution s, then s is perturbed to obtain a new solution s'. The fitness value $\boldsymbol{f}_{(s')}$ is computed for the new solution. The difference between $\boldsymbol{f}_{(s)}$ and $\boldsymbol{f}_{(s')}$ is calculated. In the *ith* internal iterations, this difference is expressed as $\Delta \boldsymbol{f}_i = \boldsymbol{f}_{(s_i')} - \boldsymbol{f}_{(s_i)}$, where L_k denotes the number of internal iterations. If $\Delta \boldsymbol{f}_i \leq 0$, s' replaces s as the new solution. Otherwise, s' is accepted with a certain probability.

$$\boldsymbol{f}(s) = 1/(d + \varepsilon) \tag{3-4}$$

where $\boldsymbol{f}$ is the fitness function, d is the calculated distance ($d = v * t$), and ε is a small constant (obstacle penalty) preventing division by zero. If an obstacle is in the path of the tower crane, a penalty is added to the objective function value.

3.2.6 Genetic Operator

The GA employs various operators to generate continuous and smooth paths. However, the random nature of path generation may result in suboptimal lengths and lack of smoothness (Li et al., 2024), impeding the effectiveness of genetic operations. To address this, we introduce two new operators: insertion and deletion. The insertion operator reconstructs and optimizes paths, while the deletion operator eliminates redundant configurations (Li et al., 2020). These operators synergistically enhance path quality and efficiency during optimization. Subsequent subsections will detail all these operators.

3.2.6.1 Selection operator

The core principle of GA ensures the survival and inheritance of optimal genes across generations (Luan and Thinh, 2023). Selection identifies superior chromosomes through three steps: First, objective function values are computed for all chromosomes. Second, fitness values are assigned based on these objective function values. A ranked-based fitness assignment method is used instead of a proportional one (Lei et al., 2013; Tuncer and Yildirim, 2012), preventing domination by a few high-performing chromosomes. Finally, chromosomes are selected based on fitness values and placed in a mating pool for new chromosome creation.

3.2.6.2 Crossover operator

Crossover is a crucial mechanism for gene reorganization within chromosomes. It combines traits from two parent chromosomes to create two offspring (Li et al., 2016). Crossover points are selected randomly (Soltani et al., 2002). Options include single-point or multi-point

crossover (Li et al., 2023). This study employs single-point crossover. The procedure involves identifying corresponding positions in two chromosomes, randomly selecting one, and crossing over at that point to generate new paths.

3.2.6.3 Mutation operator

A post-crossover random mutation was applied to all potential chromosomes. Mutation can be a random bitwise binary complement or a small gene alteration, depending on chromosome coding (Liu et al., 2023). It was uniformly applied across all the genes at a specified rate. Two genes were randomly selected, excluding the start and end configurations. The path between these genes is removed, and they are treated as adjacent points and connected using the second step of the initialization path. If a continuous path cannot be established, the process repeats with different genes until it is successful. This mutation expands the search space, enabling the exploration of distant regions from the current population. This increases the population diversity and prevents premature convergence.

3.2.6.4 Insert operator

In permutation or combinatorial optimization problems, "insert" and "deletion" operators are commonly used (Li et al., 2020; Tuncer and Yildirim, 2012). The insert operator introduces new genes or segments into the existing chromosomes, enhancing path quality and solution diversity. It selects a specific position within a path and inserts a new configuration, creating longer or more diverse routes while ensuring feasible continuous paths (Razali et al., 2023). This operator allows the GA to explore novel pathways, improve the search process, and discover optimal solutions effectively. It refines and expands the search space in path planning, thereby strengthening the GA's ability to identify superior paths.

3.2.6.5 Deletion operator

The deletion operator removes genes or segments from the existing paths, refining and simplifying them. It selects and deletes specific genes within a chromosome, resulting in shorter and potentially more efficient routes. This operator offers several benefits: it simplifies paths, reduces computational complexity (Yahia and Mohammed, 2023), and avoids convoluted or unnecessary paths, and streamlines the search process. By incorporating the deletion operator, path planning becomes more concise and relevant, enhancing the efficiency and effectiveness of the GA in finding optimal solutions (Hu et al., 2021).

3.2.7 Decision-making methods

3.2.7.1 TOPSIS method

The simulation-optimization step generates optimal solutions $\boldsymbol{S} = \{s_1, s_2, \ldots, s_n\}$, considering objective functions, constraints, and decision variables. Post-optimization evaluation of these solutions aids crane operators in path selection (Sarraf and McGuire, 2020). We focus on optimal and suboptimal paths, $\widehat{\boldsymbol{S}} = \{s_1, s_2\}$. TOPSIS, a multiple criteria decision-making analysis (MCDA) tool, ranks these paths based on multiple criteria (Madhu et al., 2023; Yavari et al., 2023). It evaluates solutions by measuring their Euclidean distance from the best and worst solutions, following the steps in Figure 3.4 and Eqs. (3-5)-(3-8) (Tansar et al., 2023).

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}} \tag{3-5}$$

$$S_i^+ = \sqrt{\sum_{j=1}^{n} \left(\left(r_{ij} - r_{ij}^+ \right) w_j \right)^2} \tag{3-6}$$

$$S_i^- = \sqrt{\sum_{j=1}^{n} \left(\left(r_{ij} - r_{ij}^- \right) w_j \right)^2} \tag{3-7}$$

$$C_i = \frac{S_i^-}{(S_i^+ + S_i^-)} \tag{3-8}$$

where, x_{ij} is the original evaluation matrix; r_{ij} is the normalized matrix; r_{ij}^+ and r_{ij}^- are best and worst solutions; S_i^+ and S_i^- are Euclidean distances from ideal and negative-ideal solutions; w_j is criteria weight; C_i is solution ranking.

3.2.7.2 Entropy weight method

Assigning appropriate criteria weights is crucial for decision-making. The entropy method, based on information theory, reduces subjectivity in weight calculation (Dong et al., 2023; Srivastava et al., 2022). It optimizes the influence of subjective factors and assigns higher weights to criteria with more information (Tansar et al., 2023). Eqs. (3-9)-(3-11) calculate these weights (Dong et al., 2023; Tansar et al., 2023).

$$w_j = \frac{(1 - e_j)}{\sum_{j=1}^{n}(1 - e_j)} \tag{3-9}$$

$$e_j = \frac{-1}{\ln(m)} \sum_{i=1}^{m} p_{ij} \ln(p_{ij}) \tag{3-10}$$

$$p_{ij} = \frac{r_{ij}}{\sum_{i=1}^{r} r_{ij}} \tag{3-11}$$

where e_j is weight entropy and p_{ij} is the normalized value of the *ith* solution for the *jth* criterion.

3.2.8 Occlusion and collision

Constantly moving machinery, tools, and workers in dynamic construction environments introduce obstacles that can disrupt crane lifting operations. These obstacles are classified as either critical or non-critical based on their potential collision risk. Critical obstacles are those located within a defined threshold on the planned paths of the crane, posing a significant risk to the lifting operation. In contrast, non-critical obstacles lie beyond this threshold and have

minimal impact on the planned path of the crane. The classification of an obstacle can change depending on its position relative to a specific path. Moreover, any movement over time may alter its influence, necessitating adjustments to the lift plan. This dynamic categorization offers a more nuanced understanding of discrete and continuous collisions on a construction site. Figure 3.5 illustrates the distinction between these types of obstacles.

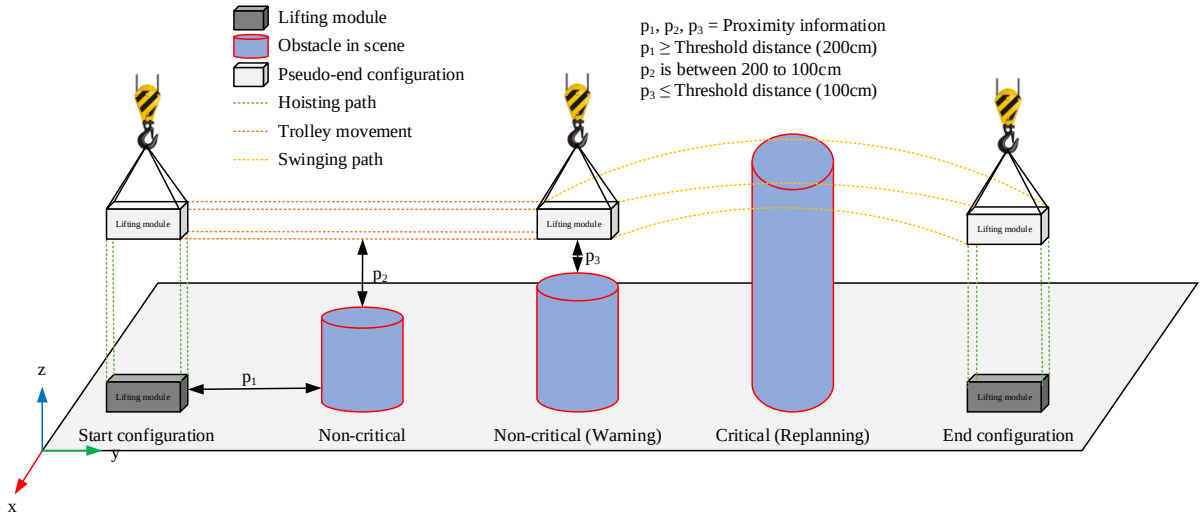


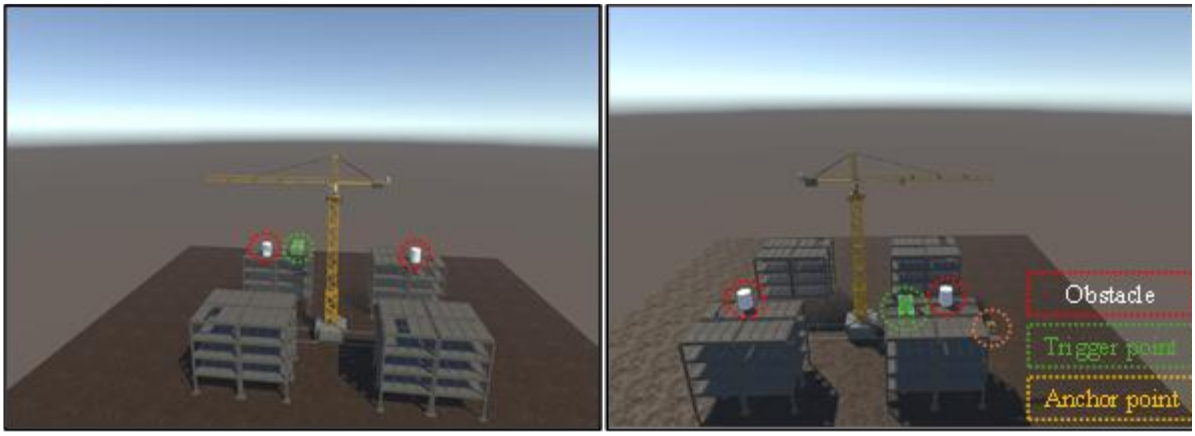
Figure 3. 5: Types of obstacles in dynamic construction environments.

3.2.8.1 User-defined obstacles scene creation

To incorporate obstacles into a scene, it is essential to define their fundamental characteristics, such as size, shape, and position. Static obstacles maintain fixed positions and orientations, whereas dynamic obstacles simulate moving entities. These obstacles are integrated as individual objects with unique attributes, controlled by user input, as shown in Figure 3.6. The plan includes creating an environment where obstacles follow predefined paths based on user inputs, with the crane occasionally moving towards them.

The goal was to develop a real-time technique to anticipate obstacle behaviors, evaluate crane motion, and determine a continuous, smooth, occlusion-free, and collision-free optimal path of the crane. Drawing from previous research (Cai et al., 2018, 2016; Dutta et al., 2020), a critical boundary was set at 100 cm. Distances exceeding 200 cm indicate a clear path, while distances

between 100-200 cm trigger a warning signal for crane operators. If obstacles approach within 100 cm, the crane automatically shuts off. These techniques enable effective planning and replanning of crane lift paths, ensuring safety and operational efficiency in dynamic construction environments.



(a) User defined obstacles from front view

(b) Used defined obstacles from back view

Figure 3. 6: User-defined obstacles: (a) Front view and (b) Back view.

3.2.9 Raycasting

Raycasting is a fundamental computer graphics and computational geometry technique for occlusion, collision, and visibility testing (Tian et al., 2018). It projects rays from a point in a specific direction to determine intersections with environmental objects (Kasapakis and Gavalas, 2017). This versatile method applies to 2D and 3D spaces across fields such as video games, robotics, and simulations (Kasapakis and Gavalas, 2017; Tian et al., 2018).

In robotics, obstacle avoidance algorithms often employ computationally expensive sampling- or optimization-based planners using volumetric maps (Nearchou and Aspragathos, 1997). For path planning and replanning in virtual environments, raycasting for robotic cranes typically implements global and local strategies (Burg et al., 2021). Global planning pre-maps the crane route by considering the entire c-space, ensuring occlusion-free and collision-free paths. Global

strategies are effective but may lack adaptability to sudden changes (Pantic et al., 2023; Paulin et al., 2021; Zhang et al., 2020). Conversely, local planning reacts in real-time to immediate obstacles, enabling quick decision-making but potentially lacking long-term foresight (Chen et al., 2023; Wei et al., 2023). Both approaches struggle to evaluate the smoothness or visibility of crane movement effectively.

A hybrid approach inspired by autonomous vehicle motion prediction techniques combines elements of local and global planning (Burg et al., 2021). It anticipates potential obstacles while employing real-time decision-making to balance safety and efficiency, enabling swift adaptation to changes and effective collision avoidance in dynamic virtual environments. Implementing this hybrid approach requires a suitable timeframe for efficient global planning while maintaining flexibility for timely changes (Halim and Ismail, 2021; Ismail et al., 2023). This approach optimizes a short timeframe and repeats the planning process for subsequent time slots. Assuming that users control obstacles through interactive inputs, their movements are predicted shortly, and crane paths are planned accordingly. Dynamic updates occur whenever the input changes. This approach results in smooth, occlusion-free, collision-free, real-time optimal paths in complex and dynamic construction environments.

3.2.9.1 Anticipating the position and trajectory of obstacles

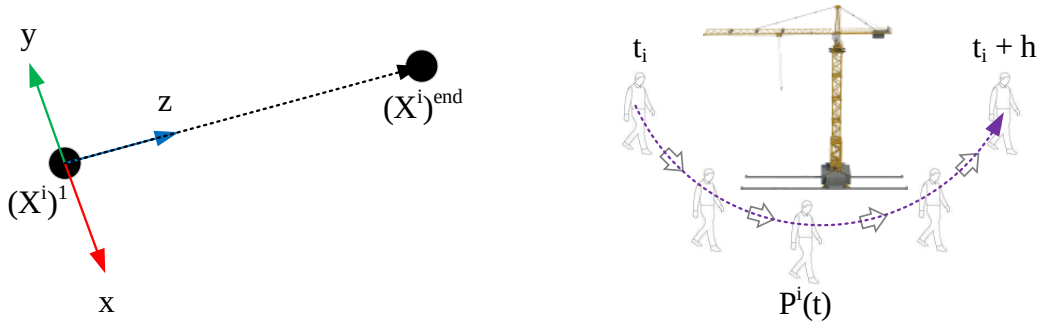
To predict the final position and trajectory of dynamic obstacles, each obstacle is defined within a time horizon, denoted as h. The start configuration $\left(X^i\right)^1$ is specified at the time t_i, and the end $\left(X^i\right)^{end}$ at $t_i + h$, as shown in Figure 3.7(a). This scenario anticipates obstacle positions within the horizon H^i, starting at t_i and lasting h seconds. The behavior of the obstacle is assumed to remain consistent over H^i.

The obstacle is modeled as a rigid body with user-specified behavior. The motion of the obstacle is simulated using the "PhysicsScene" tool in Unity over the horizon H^i, generating

predicted positions. This process produces a 3D animation curve denoted as $P^i(t)$, representing the anticipated position of the obstacle over time t, where $t \in H^i$. This curve is called the simulated behavioral trajectory, as shown in Figure 3.7(b).

Users can define viewpoints to observe obstacles with the crane in c-space, corresponding to standard angles like front, back, side, or top-down views, with a set priority order. These viewpoints from a priority queue V, each defined as 3D positions in spherical coordinates (d, θ, ϕ) in the local frame of the obstacle, where (θ, ϕ) defines the horizontal and vertical viewing angles and d is the viewing distance. The system checks viewpoints in priority order, stopping when it finds one, ensuring neither obstacle occlusion nor collision at $t_i + h$. This collision-free viewpoint becomes the goal viewpoint $(X^i)^{end}$.

The behavior of the obstacle is represented by a curve $P^i(t)$, describing expected movement over time, as depicted in Figure 3.7(b). Two crucial viewpoints $(X^i)^1$ and $(X^i)^{end}$ are identified. The system proposes generating animations between these points, denoted as $\mathbb{A}^i$. These animations are sampled and represented as cubic spline curves C^3, with each animation encoded as a_j^i, where j is the sample index. Cubic spline encoding ensures smoothness and continuity.



(a) (b)

Figure 3. 7: Obstacle behavior: (a) Animation space and local transformation, and (b) Animation curve at iteration i.

3.2.9.2 Crane motion based on animation criteria

Following the computation of the animation set $\mathbb{A}^i$ for obstacles within the time horizon H^i. One animation was selected to evaluate crane motion. The selection process evaluates all animations based on specific criteria. The primary criteria for assessing crane animation a_j^i are occlusion and collision avoidance with obstacles while providing smooth viewpoint sequences. To measure these requirements, cost measures $C_k(t)$ in the range $[0, 1]$ are proposed.

A raycasting method evaluates occlusion and collision caused by obstacles from the crane's position a_j^i(t). This method approximates obstacle geometry by sampling points $s \in [0, N]$ on the abstraction at the anticipated position $P^i(t)$. Rays are cast from the crane towards each point s, as shown in Figure 3.8. The outcome of each ray $R_s(t)$, is indicated in Eq. (3-12):

$$R_s(t) = \begin{cases} 0 & if\ visible \\ 1 & if\ occluded \\ 2 & if\ collided \end{cases} \tag{3-12}$$

Occlusion and collision are determined by the intersection of the ray with the geometry, represented here as a circle. By analyzing whether the ray intersects the front or back face, we can determine if the point $a_j^i(t)$ is inside or outside the circle. If the ray intersects the back face, the point is inside the circle and classified as a collision. If it intersects the front face of s, the point is outside the circle. If the ray does not reach a point s, it is considered occluded and, therefore, not visible. The costs for occlusion and collision are calculated using Eqs. (3-13) and (3-14) (Burg et al., 2021):

$$C_o(t) = \frac{1}{N}\sum_{s=0}^{N} \begin{cases} 1 & if\ R_s(t) = 1 \\ 0 & Otherwise \end{cases} \tag{3-13}$$

$$C_c(t) = \frac{1}{N}\sum_{s=0}^{N} \begin{cases} 1 & if\ R_s(t) = 2 \\ 0 & Otherwise \end{cases} \tag{3-14}$$

In this case, a value of $N = 20$ is used. A smooth crane motion avoids abrupt changes in distance and angle to the obstacle (Chen et al., 2023; Paulin et al., 2021; Tian et al., 2018). Two additional costs assess viewpoint changes over time: one for viewing angle and another for distance to the obstacle, evaluated at each time step δ_t. The connection between the obstacle and the crane is represented as $D_j^i(t)$, as defined in Eq. (3-15):

$$D_j^i(t) = P^i(t) - a_j^i(t) \tag{3-15}$$

View angle change is defined in Eq. (3-16):

$$C_{\Delta_{\emptyset,\theta}}(t) = \frac{\left(D_j^i(t), D_j^i(t+\delta_t\right)}{\pi} \tag{3-16}$$

A squared distance variation is proposed in Eq. (3-17):

$$\Delta d(t) = (\left\|D_j^i(t)\right\| - \left\|D_j^i(t+\delta_t)\right\|)^2 \tag{3-17}$$

A normalized cost is assigned to this distance change, as shown in Eq. (3-18):

$$C_{\Delta d}(t) = 1 - E(\Delta d(t), \lambda) \tag{3-18}$$

An exponential decay function E with $\lambda=10^{-4}$ is used. To address potential issues with this cost function, a threshold cost is introduced to encourage crane animations within a specific distance range $[d_{min}, d_{max}] = \{0, 100\}$, as indicated in Eq. (3-19):

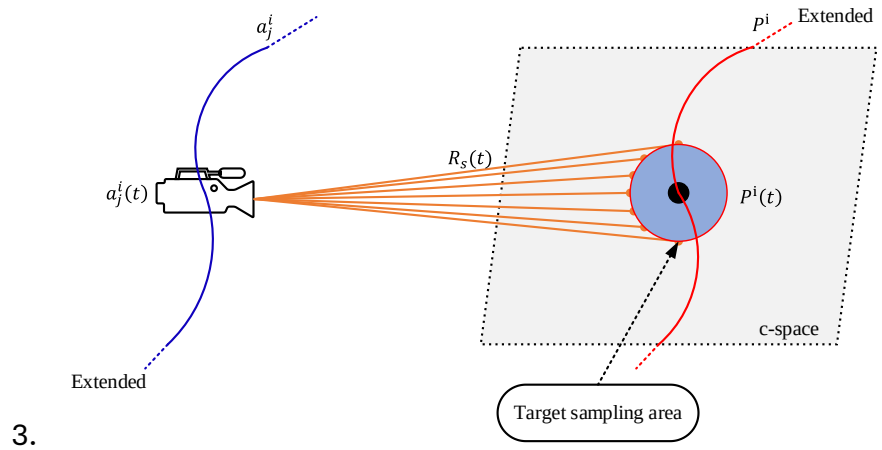
$$C_d(t) = \begin{cases} 1 & if \left\|D_j^i(t)\right\| \notin [d_{min}, d_{max}] \\ 0 & Otherwise \end{cases} \tag{3-19}$$

The overall cost of a crane animation is determined by calculating a weighted sum of individual cost criteria integrated over time. The total cost for any crane animation $a_j^i \in \mathbb{A}^i$ is computed using a discretized time integral, as shown in Eq. (3-20). The most promising crane animation a^i for time horizon H^i is based on the minimum total cost, as indicated in Eq. (3-21):

$$C = \sum_k w_k * \left[\int_{t_i}^{t_i+h} C_k(t) * G(t - t_i, \sigma) dt \right] \quad (3\text{-}20)$$

$$a^i = \underset{j}{arg\,min}\, C_j^i \quad (3\text{-}21)$$

where $w_k \in [0, 1]$ is the weight of criterion k, G is the Gaussian decay function, and σ is the standard deviation.



3.

Figure 3. 8: Ray launched towards an obstacle sampling area at time t.

3.2.9.3 Dynamic path adaptation

Two additional requirements must be addressed beyond the nominal scenario of selecting a single crane animation for a given time horizon H^i. First, the crane must be animated to track the obstacle beyond the duration h, as changes in obstacle behavior may occur due to interactive user inputs. Second, for any horizon H^i, potential collisions between the crane animation and the scene, or the obstacle occlusion, require replanning. Assuming the crane is currently animated along curve a^i, a new crane animation for the subsequent time horizon H^{i+1} is computed in two cases: (1) when the behavior of obstacle changes and (2) when considering a new obstacle behavior is deemed reasonable. In these cases, an animation is computed for an anticipated horizon of length h, but only the first steps are played to account for possible

dynamic occlusion and collision. The duration of these steps is determined by a user-defined ratio of progress along a^i. We used a horizon length $h = 5$ seconds and a progress ratio of 10%.

The new horizon generally commences at $\left(X^{i+1}\right)^1 = a^i(t_{i+1})$. If an update is required, a new crane animation is calculated for the upcoming horizon H^{i+1}, and a new viewpoint $\left(X^{i+1}\right)^1$ is selected. The transform matrix T^{i+1} is then updated to position the crane at $\mathbb{A}^{i+1}$. Subsequently, all crane animations in $\mathbb{A}^{i+1}$ are evaluated. To ensure continuity between consecutive animations, an additional cost design is introduced to favor smooth transitions. This is achieved by penalizing wide angles between the tangent vector of the current crane animation a^i, and the tangent vector of the animation at $a_j^{i+1} \in \mathbb{A}^{i+1}$, at the connection time t_{i+1}, as expressed in Eq. (3-22). The selection of the crane animation is then adjusted accordingly, as demonstrated in Eq. (3-23).

$$C_{i,i+1}(j) = \frac{\left(\dot{a}^i(t_{i+1}), \left(\dot{a}_j^{i+1}(t_{i+1})\right)\right.}{\pi} \tag{3-22}$$

$$a^{i+1} = \underset{j}{arg\ min} \left[C_j^{i+1} + w_{i,i+1} C_{i,i+1}(j)\right] \tag{3-23}$$

3.2.10 Dynamic replanning architecture

To address dynamic conditions in CALPP-RP systems, a replanning mechanism for crane lifting paths is essential. This involves determining when and how to replan. We propose a Decision Support System (DSS) and a Path Re-Planner (PRP) to meet these requirements. The replanning architecture is based on two assumptions. First, users can provide new scene information interactively, including obstacle characteristics. Second, the time delay between obstacle detection and system communication is negligible.

Figure 3.9 depicts the input-output architecture of the CALPP-RP systems. The DSS module requires inputs from planning and replanning modules. The system captures abrupt

environmental changes, considering distance and angle variations. The occlusion and collision detection engine provides proximity information and threshold distances for warning signals and replanning decisions.

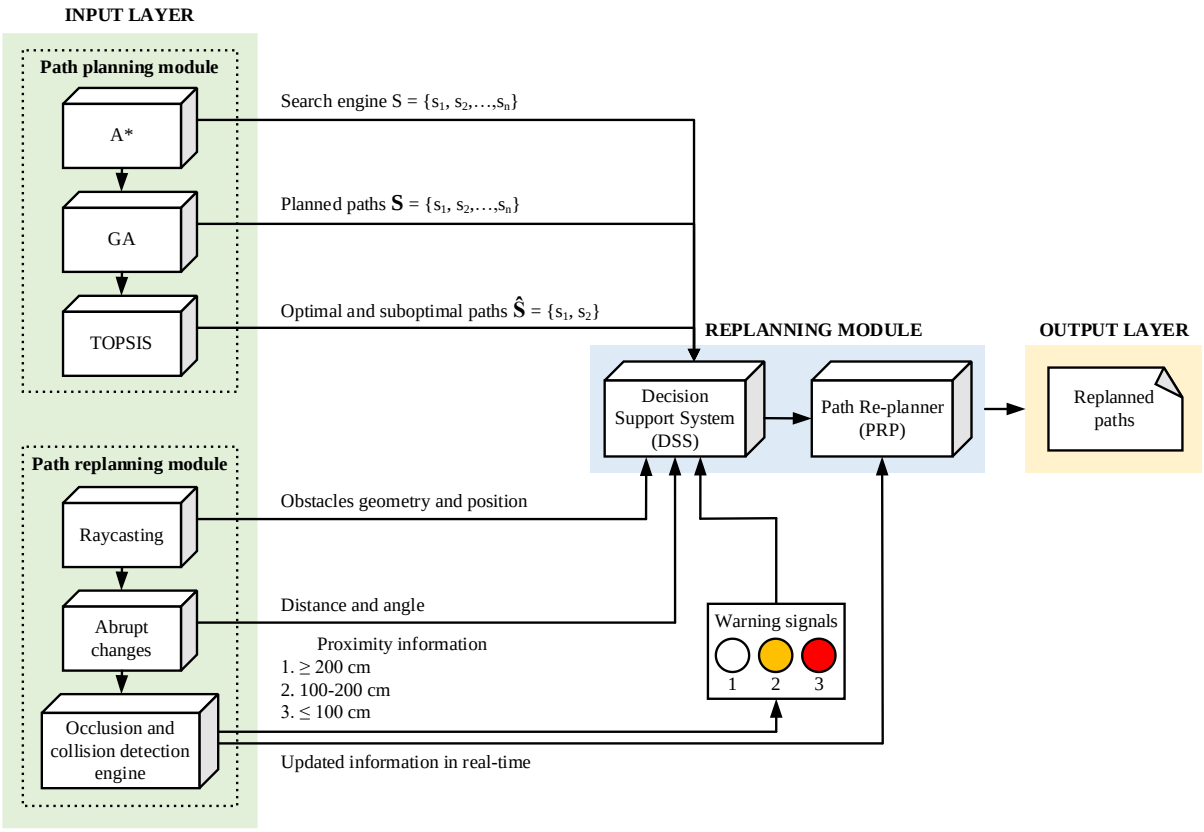


Figure 3. 9: The input-output architecture of the CALPP-RP systems.

3.2.10.1 Decision Support System

The DSS utilizes proximity data between the crane and obstacles during lifting operations. It compares crane-to-obstacle distances with predefined thresholds based on safety considerations. As outlined in Table 3.1, no signals are issued when the distance exceeds 200 cm, indicating a clear path. For distances between 100 and 200 cm, a yellow warning signal alerts operators to potential collisions, though replanning is not necessary at this stage. If the distance falls below 100 cm, a red signal triggers, indicating the need to replan the path (Dutta et al., 2020). The DSS then communicates this decision to the PRP for action.

Table 3. 1: DSS decision criteria for replanning lift paths when obstacles are detected during simulation.

Scenarios	Proximity information	Obstacles identification	Decision of replanning
1	≥200 cm	Non-critical	No
2	Between 100 and 200 cm	Non-critical	No
3	≤100 cm	Critical	Yes

3.2.10.2 Path Re-planner

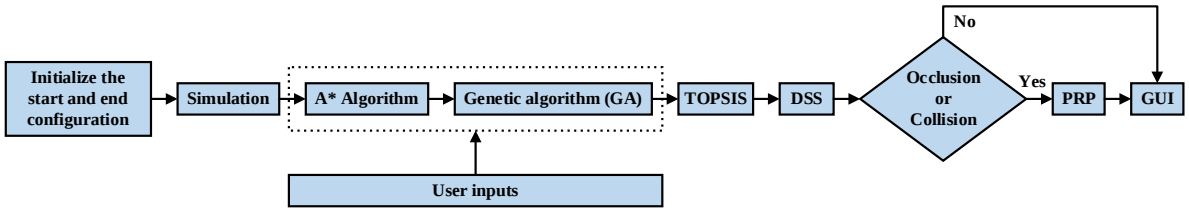
When replanning is necessary, the PRP module halts the crane and discards the remaining of the current lifting plan. It updates the task specifications with a new starting configuration for the tower crane. The system backtracks to the last valid crane configuration switch, using this as the new starting point since the stop position is invalid due to reported collisions. The end configuration remains the same as in the initial plan. When obstacles invalidate part of the planned path, the number of subsequent "genes" (path segments) to be replanned varies based on obstacle positions. The first and last operations are always included as genes to be searched. During replanning, the PRP ensures that at least two genes are explored using GA, retaining the start and end configurations on the chromosome. The PRP employs a local replanning approach, removing only invalid path candidates and replacing them with collision-free optimal paths. This method allows for efficient adaptation to obstacles while maintaining the overall structure of the lifting plan.

3.3 An integrated computing system

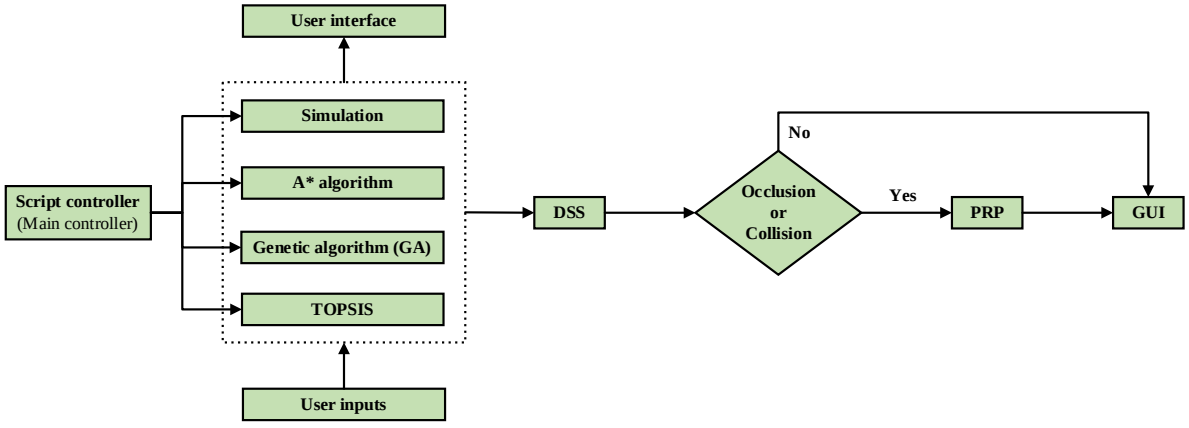
An integrated computing system combines sequential and parallel processing, creating a unified framework that leverages the strengths of both approaches. This integration parallelizes specific algorithm components while maintaining sequential structures for others, balancing simplicity and efficiency in complex path planning and replanning problems.

The primary goal of the integrated system is to optimize the computation in terms of both speed and complexity. As shown in Figure 3.10(c), the crane model was integrated with a hybrid algorithm. The crane model provided continuous inputs to the hybrid algorithm and received optimal solutions for the decision-making system. We also integrated other components into a single script that served as the main controller of our system. This controller divides the occlusion-free and collision-free crane lift path planning and replanning problems evenly across the system, running some processes in parallel (solid lines) and others sequentially (dotted lines). Crane operators, industrial practitioners, and decision-makers can monitor all planning and replanning processes in real-time. The same problem was addressed using sequential and parallel computing systems while other parameters remained constant, as shown in Figure 3.10(a) and Figure 3.10(b).

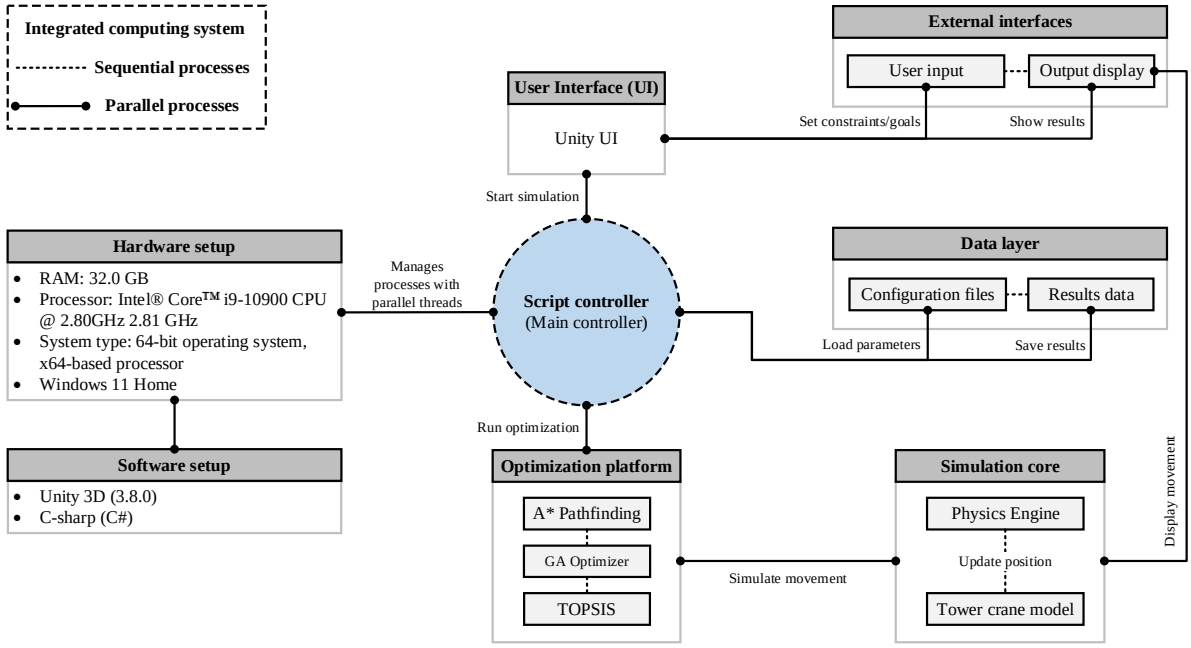
The system enhances on-site safety management by achieving occlusion-free, collision-free, and optimal paths without additional computational costs. This approach demonstrates the potential for improved efficiency and safety of crane operations through advanced computational methods.



(a) Sequential computing system



(b) Parallel computing system



(c) Integrated computing system

Figure 3. 10: Architecture of different computing systems: (a) Sequential computing system, (b) Parallel computing system, and (c) Integrated computing system.

3.4 Case study

Figure 3.11 depicts the Unity project workspace, comprising four distinct structural blocks (A, B, C, and D) and a tower crane sourced from the Unity Asset Store. A custom C# script governs crane and obstacle movement while implementing a hybrid algorithm. This script manages user input, crane positioning, rotation, and command responses. By establishing an integrated computing system, the script facilitates continuous information exchange between the crane

model and the hybrid algorithm, optimizing algorithmic outputs and enhancing system coordination.

A physics-based system was implemented to simulate authentic project site conditions, accurately replicating crane and environmental dynamics. The case study utilized a PC with an Intel® Core™ i9-10900 CPU (2.80GHz) and 32.0 GB RAM, running Microsoft Windows 11. This setup provided a robust platform for evaluating occlusion-free and collision-free crane lift path planning and replanning within the integrated computing system.

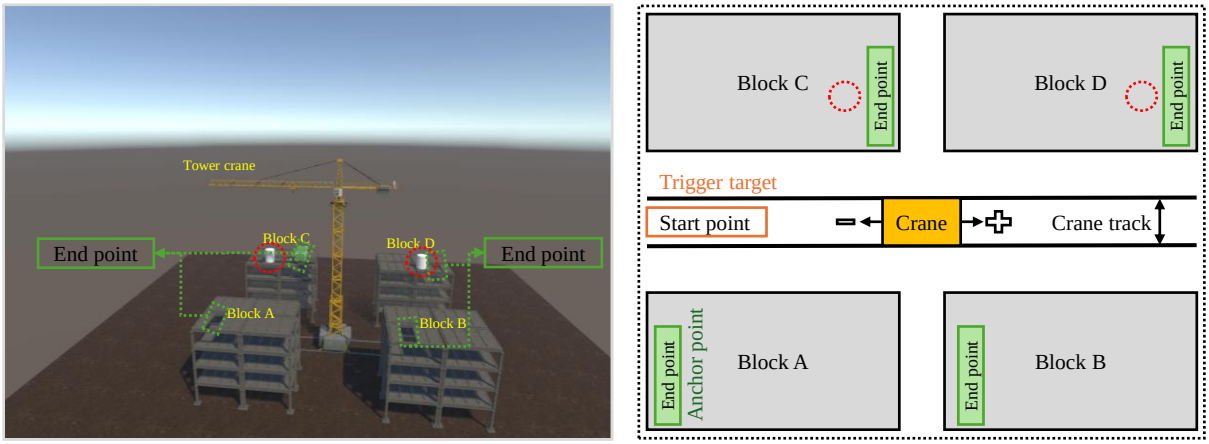


Figure 3. 11: Unity project used for the case study and scenarios selection.

3.4.1 Scenarios selection and parameter initialization

To assess the efficiency and effectiveness of the integrated computing system in path planning and replanning, this study employed four predefined scenarios encompassing diverse crane lift activities, planning processes, and replanning situations. Tables 3.2 and 3.3 list the initialized parameters for the tower crane and hybrid algorithm, respectively.

Figure 3.11 illustrates the Unity project, comprising four distinct structural blocks (A, B, C, and D). The start configuration, denoted by a yellow trigger target (Figure 3.6), was in the loading area between building blocks, following a predetermined sequence. While this start configuration remains constant across all scenarios, the final configurations vary, with each block slab containing an empty space marked in green and represented by an anchor point. The

module type (slab), weight, and dimensions remained consistent. The velocities for different DOFs were calculated as follows: $V_h = 23 * 0.8 = 18.4\, m/min$, $V_{tm} = V_{cm} = 65 * 0.8 = 52\, m/min$, and $V_{\theta_s} = V_{\theta_{lr}} = 0.8 * 0.8 = 0.64\, rpm$.

Table 3. 2: Details of the four predefined scenarios.

Scenario	Block	Floor	Module	Weight (ton)	Dimension (m*m*m)	Start configuration $(\theta_s, \theta_{lr}, d_{tm}, d_h, d_{cm})$	End configuration $(\theta_s, \theta_{lr}, d_{tm}, d_h, d_{cm})$
1	A	4F	Slab	1.44	(2*2*0.15)	(7.44, 34.08, 37.20, 5.51, -15.00)	(313.92, 32.14, 39.79, 24.37, 0)
2	B	4F	Slab	1.44	(2*2*0.15)	(7.44, 34.08, 37.20, 5.51, -15.00)	228.97, 29.72, 25.33, 237.37, 0)
3	C	4F	Slab	1.44	(2*2*0.15)	(7.44, 34.08, 37.20, 5.51, -15.00)	(65.21, 70.52, 22.16, 25.34, 0)
4	D	4F	Slab	1.44	(2*2*0.15)	(7.44, 34.08, 37.20, 5.51, -15.00)	(149.47, 32.16, 27.21, 138.76, 0)

Table 3. 3: Input parameters for the hybrid algorithm.

Parameter	value
A*	M
Population size for GA	50
Crossover rate	0.25
Mutation rate	0.75
Termination iteration	200
Termination fitness	15,000

3.4.2 Procedure

i. The workspace comprises building blocks and a tower crane imported from the Unity Asset Store, employing physics-based simulations for realistic crane lifting operations. The workspace is converted into c-space using robotic concepts (Section 3.2.3).

ii. The optimization platform integrates a hybrid algorithm combining A* and GA. C# scripts govern crane and obstacle movements and algorithm execution, establishing an

integrated computing system for continuous information exchange and decision-making (Section 3.3).

iii. Four predefined scenarios evaluate diverse crane lift activities, planning processes, and replanning situations. The objective is to lift a precast concrete slab from loading points to predetermined final positions. Empty slabs were chosen intentionally, with anchor points moved to these spaces, allowing obstacle placement along planned paths to assess avoidance capabilities.

iv. Once start and end configurations are selected in c-space, the A* algorithm explores solutions using a heuristic-guided search, incorporating an evaluation function ($f = g + h$) and bending suppression technique to generate paths $S = \{s_1, s_2, \dots, s_n\}$. This enhances GA heuristic information, accelerating convergence and reducing search space.

v. The GA optimizes path selection, iteratively selecting optimal paths from the initial population based on efficiency, safety, and other factors through a fitness function ($\boldsymbol{f}$). Optimal paths are denoted as $\boldsymbol{S} = \{s_1, s_2, \dots, s_n\}$. Table 3.4 presents the hybrid algorithm pseudocode.

vi. TOPSIS method selects optimal and suboptimal paths $\widehat{\boldsymbol{S}} = \{s_1, s_2\}$ from GA-generated paths for each scenario. Table 3.5 presents numerical results, with simulated results in Figure 3.12.

vii. User-defined obstacles in c-space represent dynamic objects, animated using curves with final positions predicted by raycasting. Proximity information and threshold values inform replanning decisions. Section 3.2.7 distinguishes critical and non-critical obstacles. Scenario 1 requires no replanning (obstacle distance ≥200 cm). Scenario 2 (100-200 cm) issues a warning signal. Scenarios 3 and 4 (≤100 cm) necessitate replanning. Table 3.6 and Figure 3.13 detail the lifting paths for each scenario.

viii. The DSS identifies critical obstacles and determines replanning necessity based on proximity. The PRP generates new optimal paths for segments affected by obstacles, avoiding occlusion and collision. Figure 3.14 displays the DSS and PRP GUI. Table 3.7 provides updated start configurations for replanning stages. The number of invalid configurations determines GA chromosome length during replanning in lifting paths.

Table 3. 4: Pseudocode for the hybrid algorithm.

Hybrid algorithm: Pseudocode	
1:	<Path> crane_lift_paths
2:	Double minFitness = PositiveInfinity
3:	List<Path> occlusion_collision_avoidance_paths
4:	List<Path> optimized_crane_lift_path
5:	Initialize start and end configuration in c-space
6:	S = GenerateInitialPopulationUsingA*($S = \{s_1, s_2, \dots, s_3\}$)
7:	**FOR** each generation in GeneticAlgorithmGenerations:
8:	**FOR** each path s in $\boldsymbol{S}$
9:	Call crane_lift_operation_strategy(s)
10:	Occlusion_collision_detected = occlusion_collision_detection(s)
11:	**IF** no occlusion or collision detected **THEN**
12:	Occlusion_collision_avoidance_paths.Add(s)
13:	**END IF**
14:	Fitness(s) = calculate_fitness(s)
15:	**IF** fitness(s) <= minFitness **THEN**
16:	minFitness = fitness(s)
17:	**END IF**
18:	**END FOR**
19:	S = SelectNextGeneration(S, fitness)
20:	Perform crossover, mutation, insertion, and deletion operations on **S**
21:	**IF** occlusion_collision_avoidance_paths is not empty **THEN**
22:	Bestpath = FindBestPath(occlusion_collision_avoidance_paths, minFitness)
23:	Optimized_crane_lift_path.add(bestPath)
24:	**END IF**
25:	**END FOR**
26:	**IF** obstacle detected (distance ≤ threshold) **THEN**
27:	Call ReplanningProcess(bestPath, Obstacle)
28:	RaycastingPrediction() to predict dynamic object positions
29:	GenerateUpdatedPath(start_config, end_config, affected_segments)

30: Update crane_lift_paths with new occlusion/collision-free path

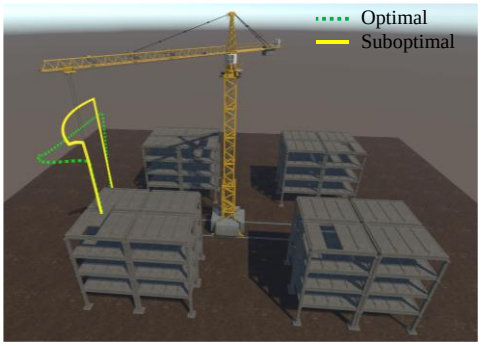
31: **END IF**

32: Apply TOPSIS method to rank and select optimal and suboptimal paths($\widehat{\boldsymbol{S}}$)

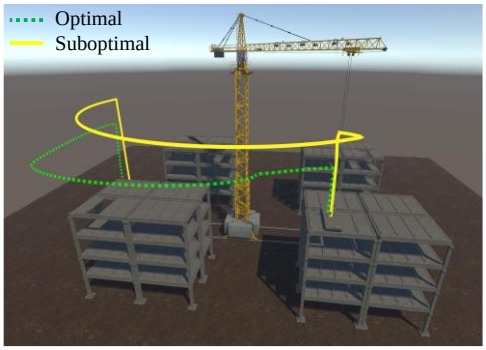
33: Execute selected path in crane operation

Table 3. 5: Optimal and suboptimal path planning for each scenario.

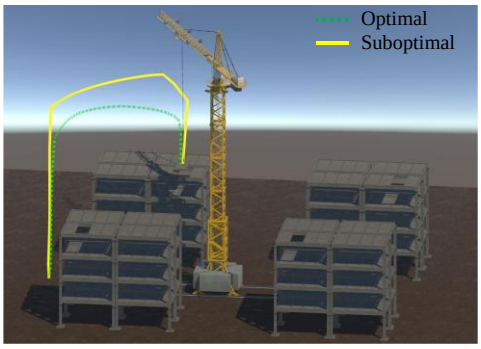
Scenarios	Block	Floor	Planned paths	Start configuration $(\theta_s, \theta_{lr}, d_{tm}, d_h, d_{cm})$	End configuration $(\theta_s, \theta_{lr}, d_{tm}, d_h, d_{cm})$	Path length	Path duration
1	A	4F	Optimal	(7.44, 34.08, 37.20, 5.51, -15.00)	(313.92, 32.14, 39.79, 24.37, 0)	109.00	25.00
			Suboptimal	//	//	120.08	31.04
2	B	4F	Optimal	(7.44, 34.08, 37.20, 5.51, -15.00)	228.97, 29.72, 25.33, 237.37, 0)	173.10	36.16
			Suboptimal	//	//	189.10	50.31
3	C	4F	Optimal	(7.44, 34.08, 37.20, 5.51, -15.00)	(65.21, 70.52, 22.16, 25.34, 0)	80.10	20.25
			Suboptimal	//	//	105.10	28.58
4	D	4F	Optimal	(7.44, 34.08, 37.20, 5.51, -15.00)	(149.47, 32.16, 27.21, 138.76, 0)	146	32.45
			Suboptimal	//	//	158.43	43.41



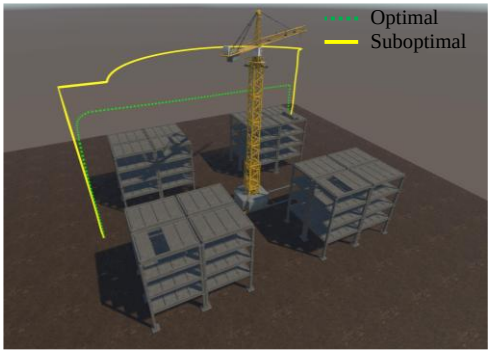
(a) Scenario 1



(b) Scenario 2



(c) Scenario 3

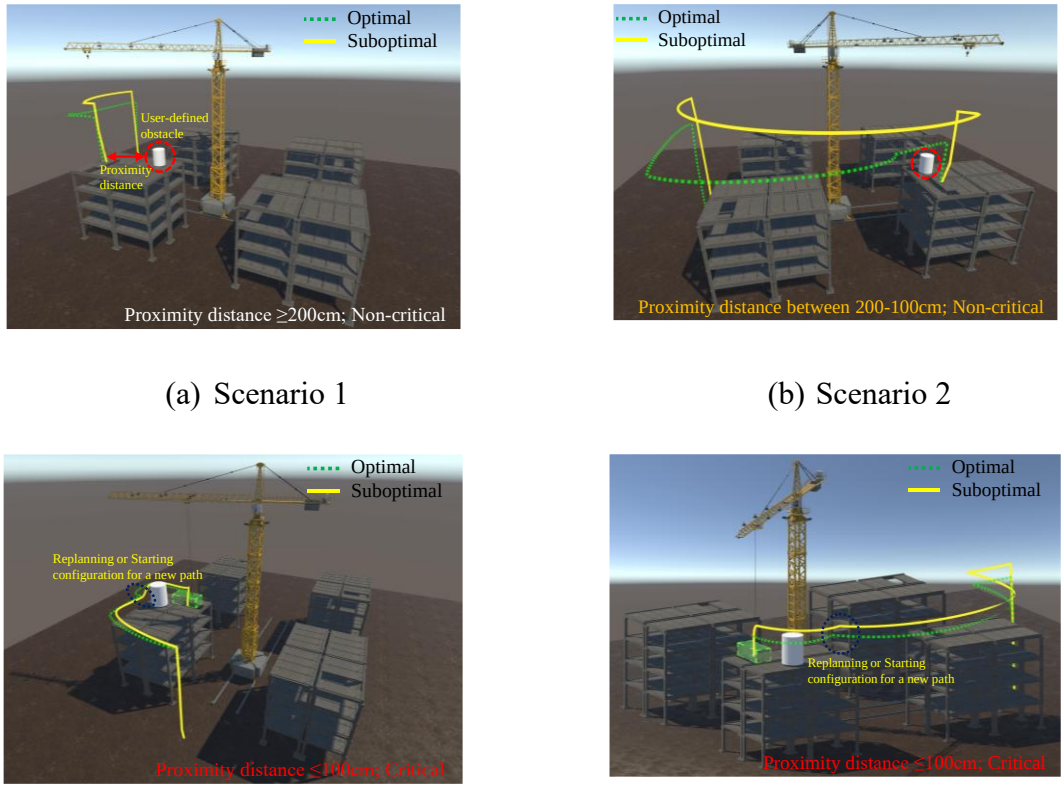


(d) Scenario 4

Figure 3. 12: Optimal and suboptimal path planning for each scenario.

Table 3. 6: Optimal and suboptimal path replanning for each scenario.

Scenarios	Proximity information	Obstacles identification	Decision of replanning	Replanned paths	Start configuration for replanning $(\theta_s, \theta_{lr}, d_{tm}, d_h, d$	End configuration $(\theta_s, \theta_{lr}, d_{tm}, d_h, d$	Path length	Path duration
1	≥200 cm	Non-critical	No	Optimal	(7.44, 34.08, 37.20, 5.51, -15.00)	(313.92, 32.14, 39.79, 24.37, 0)	109.52	25.29
				Suboptimal	//	//	120.12	31.00
2	Between 100 and 200 cm	Non-critical	No	Optimal	(7.44, 34.08, 37.20, 5.51, -15.00)	228.97, 29.72, 25.33, 237.37, 0)	173.10	36.00
				Suboptimal	//	//	189.10	50.37
3	≤100 cm	Critical	Yes	Optimal	(7.44, 34.08, 37.20, 5.51, -15.00)	(65.21, 70.52, 22.16, 25.34, 0)	92.10	20.22
				Suboptimal	//	//	112.16	28.89
4				Optimal	(7.44, 34.08, 37.20, 5.51, -15.00)	(149.47, 32.16, 27.21, 138.76, 0)	174.90	36.51
				Suboptimal	//	//	173.45	47.00



(a) Scenario 1 (b) Scenario 2

(c) Scenario 3 (d) Scenario 4

Figure 3. 13: Optimal and suboptimal path replanning for each scenario.

Table 3. 7: Replanned portions of the optimal and suboptimal paths in scenarios 3 and 4.

Scenarios	Replanned paths	New start configuration for replanning $(\theta_s, \theta_{lr}, d_{tm}, d_h, d_{cm})$	End configuration $(\theta_s, \theta_{lr}, d_{tm}, d_h, d_{cm})$	Path length	Path duration
3	Optimal	(49.56, 53.60, 16.84, 19.26, 0)	(65.21, 70.52, 22.16, 25.34, 0)	92.10	20.22
	Suboptimal	(52.17, 56.42, 17.73, 20.27, 0)	//	112.16	28.89
4	Optimal	(122.57, 26.37, 22.31, 113.78, 0)	(149.47, 32.16, 27.21, 138.76, 0)	174.90	36.51
	Suboptimal	(125.55, 27.02, 22.86, 116.56, 0)	//	173.45	47.00

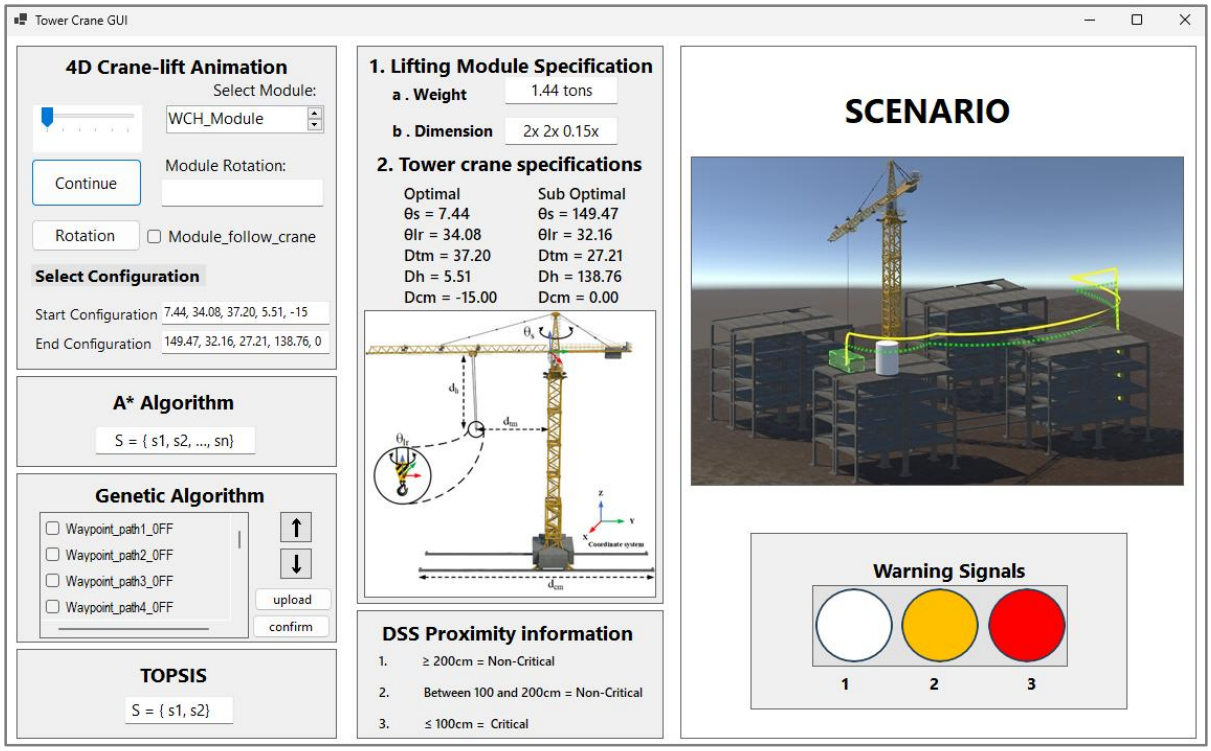


Figure 3. 14: Graphical user interface for DSS and PRP in scenario 4.

3.5 Evaluation

3.5.1 Evaluation of the simulation platform

The simulation platform was evaluated through self-assessment by the authors and semi-structured interviews with industry practitioners.

3.5.1.1 Functional testing through self-assessment

The authors conducted internal functional testing on seven key components: workspace, hybrid algorithm, visualization, animation, DSS, PRP, and dynamic data updates. They evaluated the entire workspace for crane operations to determine whether it accurately reflected the site conditions, ensuring safety, efficiency, and proper coordination. Key parameters include crane layout, accessibility, ground condition, load characteristics, operating environment, and regulatory compliance. The efficacy of the hybrid algorithm in generating optimal, occlusion-free, and collision-free paths in dynamic environments. The crane lift path visualization was assessed for accuracy, clarity, and effectiveness in operator communication. The 4D animation

enhanced understanding by adding a temporal dimension, portraying realistic movements and accurate timing. The DSS provided real-time guidance and intelligent recommendations based on environmental and crane conditions, leveraging the hybrid algorithm. Its evaluation focused on accuracy, usefulness, and impact on decision-making. The PRP assessment examined the ability of the system to dynamically replan paths in response to unexpected changes, utilizing the genetic component of the hybrid algorithm. Criteria included responsiveness, efficiency, and quality of alternative paths. Finally, the evaluation scrutinized the capacity of the system to integrate real-time updates of environmental data and crane specifications, ensuring the hybrid algorithm and DSS maintained up-to-date path accuracy.

3.5.1.2 Results of semi-structured interviews

Following a self-assessment of the functional testing, semi-structured interviews were conducted with five operators, each boasting nearly two decades of experience. The aim was to evaluate the merits and limitations of the proposed integrated computing system for occlusion-free and collision-free crane lift path planning and replanning in dynamic environments. The interviewees offered positive feedback, emphasizing: (1) real-time planning and replanning capabilities, (2) practical and manipulable visualized crane paths, (3) immersive animation, (4) realistic occlusion and collision detection, (5) real-time communication between DSS and PRP systems, (6) effective handling of complex, high-dimensional environments, and (7) dynamic data flow.

The operators also identified areas for improvement, suggesting: (1) development of a user-friendly GUI for stakeholders, (2) implementation of more advanced algorithms, (3) minimization of manual steps for full automation, (4) consideration of multiple crane operations, and (5) adjustment of threshold values for more realistic lifting module movement.

In response to this feedback, we have successfully addressed the GUI design, manual processes, and threshold value issues, significantly enhancing user experience and streamlining operations. We acknowledge the recommendations regarding advanced algorithms and multiple crane integration, and we are committed to exploring these aspects in future research to refine our methodology further. The validation from experienced practitioners underscores the effectiveness of our approach in dynamic environments.

3.5.2 Evaluation of optimization platform

The performance of the hybrid algorithm was evaluated against six other algorithms, including Dijkstra, A*, GA, PGA, GA-SA, and PSO-SA, using five criteria: computational time, convergence curve, success rate, solution stability, and path smoothness. All algorithms were tested within an integrated computing system under identical parameters, including c-space size, number of DOFs, tower crane operational strategy, iteration count, and fitness values. This ensures a fair and consistent comparison. The analysis was conducted from 30 independent runs of each algorithm, rigorously examining the results across the specified criteria.

3.5.2.1 Computational time

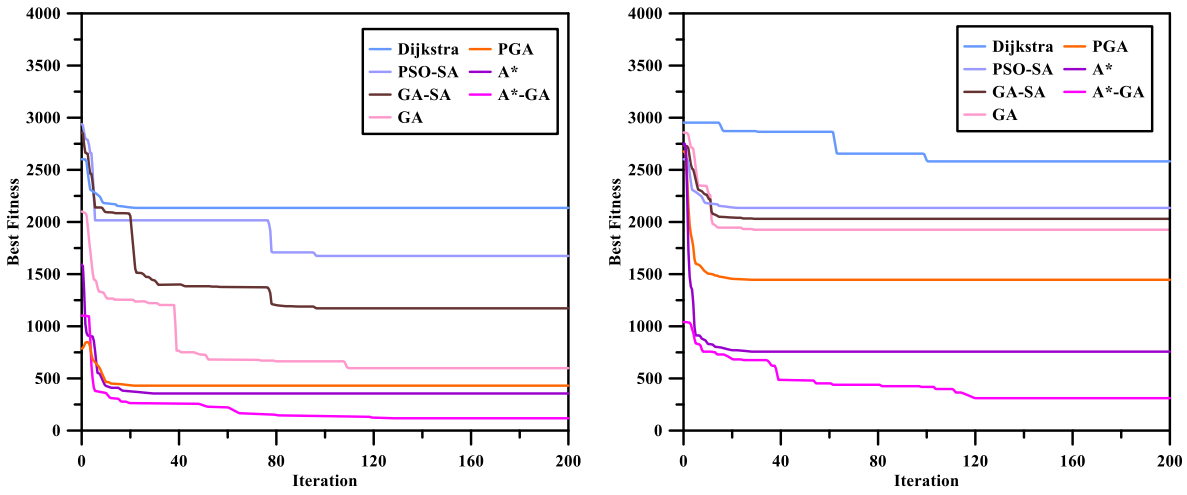
Algorithm computational time within an integrated computing system is influenced by multiple factors, including hardware performance, platform, iteration count, population size, c-space dimensions, DOF count, environmental changes, and occlusion and collision checks. Our evaluation maintained consistent conditions across all algorithms, varying only the algorithm under comparison. Table 3.8 displays the average computational times for four scenarios. The proposed hybrid algorithm achieved an average of 0.42 seconds, outperforming A* (4.64s), GA-SA (5.70s), PGA (6.93s), GA (11.37s), Dijkstra (13.65s), and PSO-SA (40.58s). Notably, these computational times are lower than those reported in previous studies, attributable to an integrated computing system (AlBahnassi and Hammad, 2012; Hu and Fang, 2016; Tuncer and Yildirim, 2012; Vu et al., 2022; Zhu et al., 2022).

Table 3. 8: Average computational time of various algorithms across four scenarios.

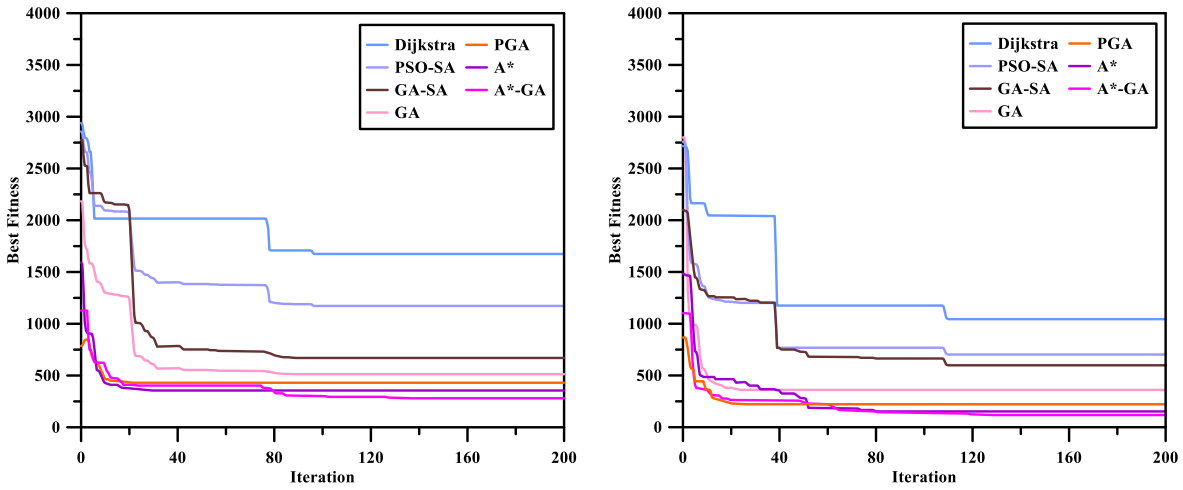
	A*-GA	A*	GA-SA	PGA	GA	Dijkstra	PSO-SA
Scenario 1	0.22	4.31	5.41	6.69	10.80	12.95	39.94
Scenario 2	0.50	4.98	5.81	6.94	11.59	13.91	40.94
Scenario 3	0.38	4.59	5.71	6.87	11.43	13.72	40.67
Scenario 4	0.59	4.66	5.85	7.23	11.66	14.00	40.75
Mean	0.42	4.64	5.70	6.93	11.37	13.65	40.58

3.5.2.2 Convergence curve

Convergence performance is a crucial indicator of the efficacy of an optimization platform. Figure 3.15 illustrates the convergence curves of various algorithms across four scenarios, demonstrating how optimization performance improves over iterations. The convergence rate signifies the speed of the algorithm in reaching a solution. A*-GA distinguishes itself through its combination of diverse exploration, environmental adaptability, and effective genetic operators. It excels in computational efficiency and handling complex environments. By adeptly navigating search spaces with A* while harnessing the exploration power and adaptability of genetic algorithms, A*-GA surpasses other algorithms. This hybrid approach yields a superior convergence curve, particularly suitable for dynamic environments in crane lift path planning where occlusion and collision avoidance are paramount.



(a) Scenario 1 (b) Scenario 2



(c) Scenario 3 (d) Scenario 4

Figure 3. 15: Convergence curves of various algorithms across four scenarios.

3.5.2.3 Success rate

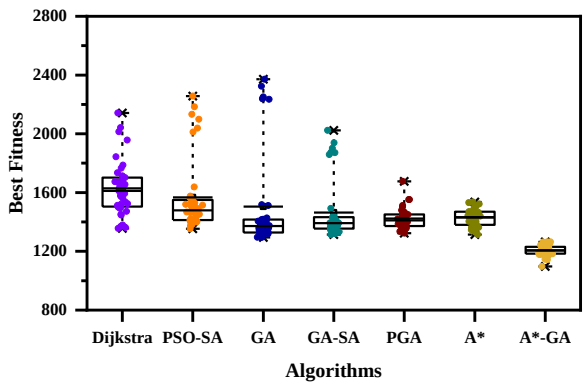
The efficacy of the algorithm is primarily determined by its ability to generate occlusion-free and collision-free optimal and suboptimal paths, resulting in high pathfinding success rates. Success rates are influenced by environmental complexity, occlusion and collision detection efficiency, and adaptability to dynamic changes. Our evaluation employed a real-time raycasting technique for efficient detection, effectively adapting to dynamic environments. Table 3.9 presents average success rates, with the proposed hybrid algorithm achieving the highest mean (99.95%) and Dijkstra the lowest (92.46%). These success rates align with those reported in previous studies, underscoring the competitive performance of the algorithm (Ali et al., 2005; Kasapakis and Gavalas, 2017; Tian et al., 2018; Zhu et al., 2022).

Table 3. 9: Average success rate of various algorithms across four scenarios.

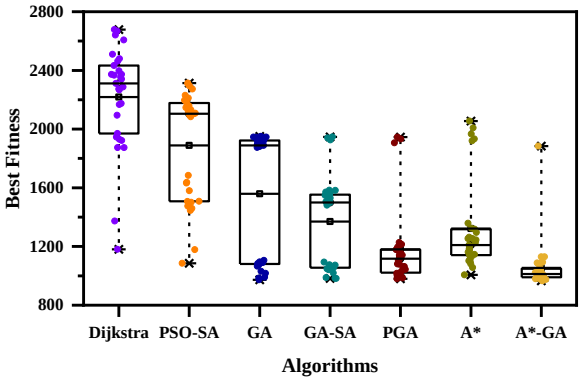
	A*-GA	A*	PGA	GA	GA-SA	PSO-SA	Dijkstra
Scenario 1	100%	100%	100%	100%	100%	100%	100%
Scenario 2	100%	100%	100%	100%	100%	100%	100%
Scenario 3	99.89%	98.54%	97.68%	96.67%	96.51%	95.11%	89.90%
Scenario 4	99.91%	99.21%	98.37%	97.82%	97.70%	96.54%	79.92%
Mean	99.95%	99.44%	99.01%	98.62%	98.55%	97.91%	92.46%

3.5.2.4 Solution stability

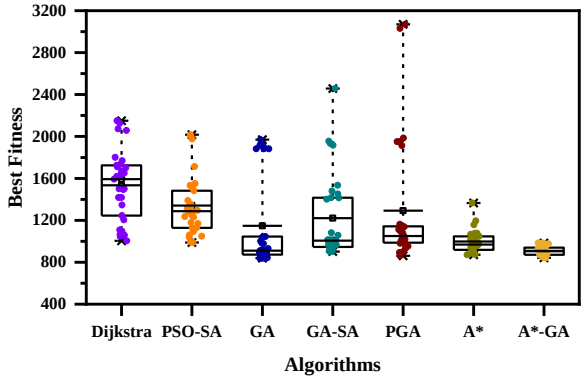
Figure 3.16 illustrates the solution stability of various algorithms across four scenarios. Solution stability denotes the robustness and reliability of solutions over time, particularly in dynamic environments. A*-GA demonstrates superior performance, evidenced by lower means and medians in the box plots and a smaller interquartile range (IQR) than other algorithms. This enhanced stability stems from adaptability to dynamic environments, diverse solution space exploration, local optima avoidance, efficient replanning, and effective genetic operators. The efficiency of the hybrid algorithm, coupled with the integrated computing system, facilitates faster computation, enabling real-time replanning and swift adaptation to environmental changes.

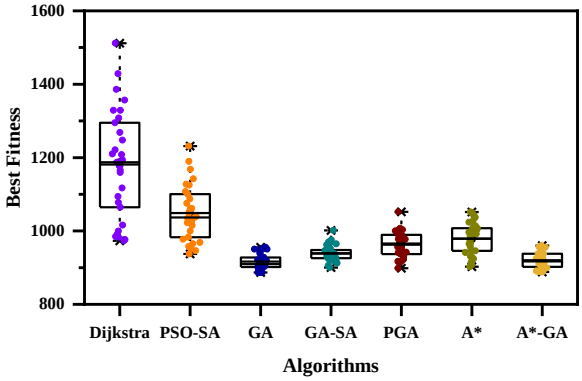


(a) Scenario 1



(b) Scenario 2





(c) Scenario 3 (d) Scenario 4

Figure 3. 16: Solution stability of various algorithms across four scenarios.

3.5.2.5 Path smoothness

Two strategies were implemented to achieve path smoothness. First, bending suppression was introduced to the A* algorithm, while insertion and deletion operators were incorporated into the GA to minimize abrupt changes. Second, constraint criteria (Eqs. 3-15-13-9) were integrated to address sudden angle and distance variations. Paths with angle changes below 5° or distance changes below 100 cm incurred penalties. These constraints, embedded in the fitness functions, yielded smoother, more desirable paths. Table 3.10 presents a comparative analysis of redundant paths with and without path constraints. Table 3.11 displays the mean number of redundant points without constraints. The results demonstrate the effectiveness of the proposed constraints in eliminating unfavorable configurations across multiple methods.

Table 3. 10: Path redundancies with constraints (C) and without constraints (NC) observed over 30 iterations.

	A*-GA		A*		PGA		GA		GA-SA		PSO-SA		Dijkstra	
	NC	C	NC	C	NC	C	NC	C	NC	C	NC	C	NC	C
Scenario 1	30	0	30	0	30	0	28	7	30	0	30	0	20	10
Scenario 2	30	0	30	0	29	0	24	12	29	0	30	0	18	12
Scenario 3	30	0	29	0	29	0	27	5	30	0	30	0	22	8
Scenario 4	30	0	30	0	30	0	29	1	30	0	29	0	19	11

Table 3. 11: The average number of redundant points per path without constraints.

	A*-GA	A*	PGA	GA	GA-SA	PSO-SA	Dijkstra
Scenario 1	1.25	2.00	1.60	1.35	1.50	1.25	2.75
Scenario 2	1.60	1.90	1.50	1.25	1.55	1.35	2.65
Scenario 3	1.40	2.10	1.75	1.45	1.35	1.20	2.90
Scenario 4	1.35	1.75	1.55	1.20	1.45	1.30	2.70

3.6 Discussion

In this section, we examine the computational times associated with sequential, parallel, and integrated computing systems, the principles underlying each system, and the functionality of hybrid algorithms. In sequential computing, we typically iterate through multiple generations, assessing each candidate solution within a population. In contrast, parallel computing distributes the workload evenly across processors, enhancing efficiency. Integrated computing systems offer optimizations that extend beyond basic parallelization, incorporating improved memory management, expedited inter-process communication, and hardware acceleration. The computational time for each system is determined using Eqs. (3-24)-(3-26), and the percentage difference is calculated as outlined in Eq. (3-27):

$$T_{seq} = N * M * T_{eval} \tag{3-24}$$

$$T_{parallel} = \frac{T_{seq}}{P} \tag{3-25}$$

$$T_{integrated} = \frac{T_{parallel}}{E} \tag{3-26}$$

$$Difference(\%) = \left(\frac{Orignal\ time - Time\ difference}{Orignal\ time}\right) * 100 \tag{3-27}$$

Here, N represents the number of generations, M denotes the population size, T_{eval} is the time required to evaluate a single candidate solution, P is the number of processors (Set at 4), and E is the performance gain factor (set at 1.5).

Table 3. 12: Comparison between sequential, parallel, and integrated computing systems.

System	Computational time(s)
Sequential computing system	0.77
Parallel computing system	0.56
Integrated computing system	0.42

3.7 Chapter Summary

The chapter presents an integrated computing system for collision-free and occlusion-free path planning and replanning of robotized cranes in dynamic construction environments. Traditional sequential and parallel computing systems face limitations such as computational inefficiency, synchronization issues, and reduced adaptability in high-dimensional settings. To address these challenges, the proposed system combines sequential and parallel processing in a unified framework, utilizing a hybrid A*-Genetic Algorithm (GA).

The hybrid algorithm integrates A* for heuristic-guided path exploration and GA for optimizing paths through operators like crossover, mutation, insertion, and deletion. The system employs raycasting for efficient collision and occlusion detection, avoiding computationally intensive volumetric mapping. Decision-making is further enhanced using the TOPSIS method to evaluate and rank optimal and suboptimal paths. A Decision Support System (DSS) ensures occlusion-free planning, while a Path Re-Planner (PRP) dynamically adjusts paths in response to environmental changes.

The system's capabilities were demonstrated through simulations in Unity 3D, incorporating realistic crane operations and dynamic environments. The results revealed that the integrated computing system outperformed traditional sequential and parallel systems, achieving an 83.33% and 33.32% reduction in computational time, respectively. The hybrid algorithm demonstrated superior convergence rates, success rates, and solution stability while ensuring smooth and efficient paths.

By addressing critical challenges such as computational latency, path smoothness, and real-time adaptability, the proposed system offers significant advancements in crane lift path planning and safety management. Its contributions include a novel hybrid algorithm, a physics-based simulation platform, and a comprehensive framework for real-time decision-making.

This research has practical implications for enhancing safety, efficiency, and automation in construction, with potential applications across various industrial domains. Future work will focus on further automation and scalability to multi-crane operations.

CHAPTER 04

PREDICTING DEGRADED LIFTING CAPACITY OF AGING TOWER CRANES: A DIGITAL TWIN-DRIVEN APPROACH

4.1 Introduction

Tower cranes are essential in modern construction projects, enabling the lifting and moving of heavy materials, equipment, and tools to precise locations on construction sites (Code of Practice for Safe Use of Tower Cranes). Over prolonged periods of use, however, aging tower cranes may experience structural fatigue, corrosion, and deterioration due to exposure to high-cycle and variable-amplitude loads, harsh working environments, and improper operation and maintenance, leading to degraded lifting capacity (LC) (Das et al., 2018; Gu et al., 2021; Guo et al., 2021; Kulka et al., 2018; Pal et al., 2018; Shin, 2015; Vukelic et al., 2019). Such degradation can increase the risk of accidents and jeopardize the safety of workers and nearby properties, resulting in severe injuries or fatalities (Lee et al., 2020). For instance, in 2022, the tower crane at the "Sau Mau Ping" construction site in Hong Kong collapsed and killed three workers, where the operator set the manufacturer-specified load without considering the deteriorated load and/or capacity (Steven Chun-yin et al., 2022). Similar incidents have occurred worldwide, including a crane collapse that killed 18 people in 2016 in Dongguan, China (Lan et al., 2017), and another crane collapse that killed seven people and injured dozens more in 2008 in New York City in the United States (Linda Levine, 2008). Thus, there is still a practical need for real-time safety monitoring systems to detect and prevent accidents caused by aging tower cranes with degraded LC.

Various safety monitoring systems have been developed and adopted in the construction industry to address crane safety, including load moment indicators (LMIs) (Fang and Cho, 2017; Kalairassan et al., 2017; Neitzel et al., 2001; Shaikh and Kumar, 2016), anti-two-blocking (ATB) systems (Rayco Wylie, 2012; Walbridge et al., 2022), and crane cameras (Chen

et al., 2023; Elgendi et al., 2023). LMIs measure the load on the crane and provide a warning when the load is near or exceeds the crane's maximum LC (Fang and Cho, 2017). Anti-two-block systems prevent the crane from collapsing by limiting its operation when the hook block is too close to the boom tip (Walbridge et al., 2022). Crane cameras monitor the crane's movement and detect unsafe behavior by recording a visual perspective (Chen et al., 2023). However, while these systems can improve safety on construction sites, they are generally not designed to address the specific issue of aging tower cranes with degraded LC. For instance, LMIs, ATB systems, and crane cameras do not consider the effects of wear and tear on crane components or the health level of the crane and, thus, fail to predict the internal degradation in lifting capacities, which can result in unexpected accidents due to unintentional over-loading. Furthermore, due to the unavailability of historical data on crane operations, the complexity of the deterioration mechanism itself, and the unpredictability of external deterioration-related factors, it remains an open question of how to estimate and update the aging-induced lifting capacity degradation.

Many studies have been conducted to tackle the issues of data availability and deterioration mechanism in understanding the degraded LC of tower cranes, which can be categorized into two groups: (1) data-driven methods (DDM) and (2) model-based methods (MBM). DDM uses real-time, historical, and maintenance data (Roman et al., 2019) to predict deterioration patterns (Roman et al., 2021). Several studies have used DDM to predict performance degradation in tower cranes (Tran et al., 2012). DDM is only effective when data are abundant, and without proper domain knowledge of the deterioration mechanisms, it may be less reliable when applied practically. On the other hand, MBM requires an understanding of failure mechanisms such as cracks, wear, deflection, friction, and stress. One can estimate performance degradation by developing a mathematical model that quantitatively describes the crane's degradation. Extensive experiments are required for parameter identification, varying from case to case. Due

to the complex nature of tower cranes, MBM may be less efficient. To overcome the limitations of DDM and MBM, Liu et al. (2021) proposed a hybrid approach that combines both methods (DDM-MBM). This approach collects data from multiple sensors since a single sensor cannot capture all the relevant parameters that affect degradation. The use of multiple sensors may increase complexity, but it results in more accurate predictions. Additionally, the mechanical model used in MBM can simulate the structural performance of the tower crane, which increases the interpretability of DDM. Therefore, a hybrid approach that integrates multi-sensor data with an interpretable model could be a possible solution to enable real-time monitoring and prediction of tower cranes' safety conditions considering degraded LC.

As an instance of hybrid DDM-MBM methods, the digital twin-driven (DTD) approach creates a replica of a physical system, aiming to enable real-time data acquisition, response prediction, and human-machine interaction (Jiang et al., 2022). Numerous studies have adopted DTD approaches in predictive maintenance in manufacturing machines, autonomous management and control systems, product design, diesel engine traceability and control, machining process observation and evaluation, and tower crane stability and hoisting safety analysis (Cheng et al., 2020; Deebak and Al-Turjman, 2021; Jia et al., 2023; Kušić et al., 2023; Liu et al., 2021a, 2021b; Liu et al., 2023; Shen et al., 2023; Wu et al., 2021; Wu et al., 2022; Zhu et al., 2021). However, despite the abundant application of DT in the abovementioned fields, research is limited in exploring the prospective role of the DTD approach in predicting the deterioration of a complex machine and/or system in the construction industry with real-time data. Still, the examples in adjacent fields illustrate the flexibility and potential of the DTD approach in improving efficiency, safety, and decision-making in various applications.

Aiming to address the real-time safety monitoring and alerting of aging tower cranes, this study develops a DTD framework and model to predict degraded LC after different usage periods. The framework integrates theoretical and numerical analyses of tower crane fatigue and

degradation behavior with real-time data collected from a physical tower crane. Machine learning models, including Random Forest (RF) and Support Vector Machine (SVM), are employed for feature selection and LC prediction. Specifically, a scaled-down tower crane prototype is adopted as a demonstrative case to illustrate the feasibility and effectiveness of the DTD framework. The predicted degraded load charts of the prototype for each decade of usage from 0 to 70 years are presented for validation purposes. The proposed DTD framework allows for continuous monitoring of the crane's safety performance considering the degradation in LC with real-time data and numerical model, which is expected to significantly improve safety in construction sites with aging tower cranes.

4.2 Methodology

4.2.1 Digital twin-driven framework for degradation in machines

The primary goal of this study is to use a DTD approach to predict degradation in machines. A DTD framework is formulated based on the mapping between physical and virtual machines, as illustrated in Figure 4.1. This framework aims to: (1) provide background information; (2) highlight issues; (3) define data requirements; (4) simplify mapping problems between virtual-physical spaces; (5) establish two-way communications between all parts; and (6) make decisions and feedback for improvement. A DTD approach for degradation must have five parts: physical space, virtual space, IoT-based connection, real-time database, and control system (Jiang et al., 2023; Kong et al., 2023; Pauwels et al., 2023). Machine degradation is predicted by mapping or fusing the DTD components using DT data.

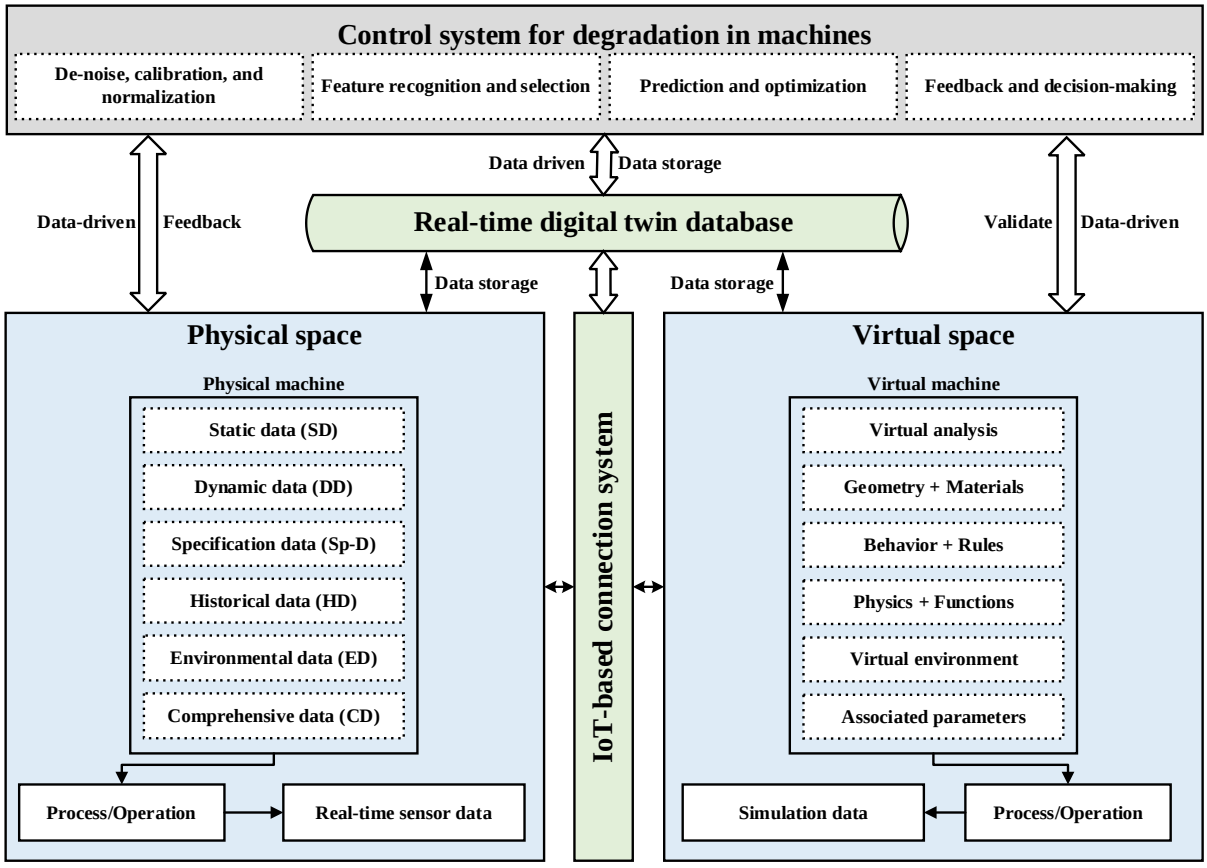


Figure 4. 1: A DTD framework for monitoring degradation in machines.

4.2.1.1 Physical space

The physical space includes all physical activities, operational processes, and work environments that induce degradation (Jiang et al., 2022). Data covering element composition, dynamics, failure mechanisms, and deterioration factors must be collected to advance the DTD model's accuracy and efficacy (Cheng et al., 2020). Sensors and transducers having networking capability via Wi-Fi or ethernet should be connected to a physical machine to collect multi-source, heterogeneous, and real-time data. Aside from that, the physical machine has a built-in data logging system that records usage information (He et al., 2021). Additional data is required to enhance the DTD model's quality, such as static data (SD), specification data (Sp-D), historical data (HD), environmental data (ED), and comprehensive data (CD), which will be explained in Section 4.2.3.3. Once we collect all the required data from the physical and virtual

spaces and build a two-way connection, decisions will be made. Feedback will be provided for machine monitoring, control, safety, and improvement.

4.2.1.2 Virtual space

A DTD approach requires a CAD model or virtual body that accurately reflects the physical machine's features, geometry, materials, behaviors, physics, rules, functions, and environment (Christiand, 2020; Zhao et al., 2019). The data fidelity is determined by the level of detail (LOD) and varies from case to case (Jinfeng Liu et al., 2019; Zhang et al., 2021). Due to data fidelity from the physical space, the static virtual body becomes "alive." It can mirror its element composition, dynamics, and status in the virtual environment to display the real-time changes of the physical machine operation and monitoring. Its primary purpose is to realistically describe and dynamically simulate physical entities and processes in multi-dimensional and multi-time scale ways (Jiang et al., 2022). The virtual body offers an open interface for data-driven modeling, optimization, and the capacity to predict the physical machine's current status and RUL (He et al., 2021).

4.2.1.3 IoT-based connection

According to ISO/OSI communication standards, the communication system for the DTD model can be wired or wireless (Unit II ISO/OSI Model in Communication Networks). An IoT-based connection is utilized for real-time data transfer and parsing. This link is supposed to provide two-way communication between physical and virtual spaces. Furthermore, all components of the DTD model should be fused and mapped to detect/resist the real-time changes in a physical space and feedback from the control system (Cao et al., 2022; Wu et al., 2021). It is also considered the secure connection and data transmission from the physical machine to the virtual machine and vice versa.

4.2.1.4 Real-time digital twin database

A multi-dimensional and multi-scale virtual simulation model is integrated with a physical model and a control system via real-time DT data. Data is the DTD model's fundamental concept, allowing the virtual model to run and make decisions. It is primarily answerable for interpretation, classification, storage, pre-processing, maintenance, and data testing (Deebak and Al-Turjman, 2021).

4.2.1.5 Control system

The control system regulates and predicts degradation based on real-time DT data. It is the hybrid mode and the most fundamental phase in a DTD model. A control system is used to analyze, and visualize real-time changes, make decisions, and provide feedback for improvement (Zhang et al., 2021). It consists of (1) data pre-processing and (2) a functional module. Pre-processing is mainly responsible for online monitoring, de-noise, calibration, normalization, analysis, statistics, and evaluating raw and real-time data collected from physical and virtual spaces (Cheng et al., 2020). This module also includes feature recognition and selection, which induces degradation (Priyanka and Thangavel, 2022). In addition, the functional module incorporates both prediction and visual display (He et al., 2021). The predicted results are used to make decisions and improve the physical machine with feedback. The operator or decision-making team directly implements the optimization and adjustment plan.

4.2.2 Fatigue behavior and degradation analysis

Regardless of the appropriate design, a tower crane is considered a fatigue-prone structure that can collapse in service (Liu et al., 2021; Mdot,2016). Based on the S-N curve (Fatigue curve, Wohler curve), a tower crane is classified as a high-cycle and super-high-cycle fatigue structure (Bandara et al., 2015). The fatigue curve is the relationship between the amplitude of cyclic

stress and the number of cycles till failure. Fatigue behavior can be analyzed using stress-life, strain-life, non-destructive, fracture mechanics, and damage mechanics.

The stress-life approach entails plotting stress against the number of cycles until failure (Dong et al., 2021). This approach needs extensive experiments, and a single stress-life plot has limitations in predicting plasticity, deformation, and mean stress. Similarly, the strain-life method includes plotting strain against the number of cycles till failure (Jiang et al., 2021). This approach requires the stress-strain curve and plasticity nucleation to determine the life of the transition between the plastic and elastic regions under fatigue loading. So, it is only applicable when the crane structure's crack initiation life is required. Fatigue can be studied non-destructively by measuring and plotting the values of inherent properties such as cracks, decreased strength, and degraded stiffness (Six et al., 2020). However, it requires a detailed understanding of the parameter's chemistry, and it is challenging to cover all parameters (Liu et al., 2021). Furthermore, in fracture mechanics, flaws and micro-cracks are pre-existing defects (Taheri et al., 2013). The degradation rate is determined using a crack growth law and a stress intensity factor as a function of crack depth, shape factor, and stress range, but we don't consider crane sections with pre-existing defects. Finally, damage mechanics require adequate assessment models, laws, and patterns to characterize damages, residual stresses, and strain localization (Bandara et al., 2015; Fatemi and Vangt, 1998; Taheri et al., 2013; Teng et al., 2020). Damage mechanic is appropriate for addressing tower crane degradation because it covers optimum factors that directly and indirectly influence crane degradation. For example, micro-plastic strain is often overlooked in low-cycle problems. Still, it must be considered when high-cycle or super-high-cycle fatigue damage occurs in the elastic range, like in a crane. Macro-plastic strain may arise because of this scenario, leading to structural deterioration under long-term operational loading.

Steel has infinite design life if the stress range is low. However, the crane is a fatigue-prone structure, and its components experience a spectrum of stress while operating. For instance, a tower crane has a higher stress range when the hook raises a hefty load at full stretch and lifting radius. Aside from that, crane loads are cyclic and have varied amplitudes. Elastic deformation occurs in a crane structure, develops stress, and accumulates residual stresses over time. When residual stresses reach a threshold value, cracks are formed, and the crane will collapse. Therefore, we must measure the loads and the accompanying deflection/deformation/vibration from the physical tower crane to predict degradation.

4.2.2.1 Fatigue damage accumulation analysis

There are two basic approaches for assessing fatigue damage: (1) constant amplitude fatigue loading and (2) variable amplitude fatigue loading. The crane load is stochastic cyclic. Therefore, fatigue damage accumulation varies with loading. The linear cumulative damage criterion does not apply in the tower crane scenario (Chaboche and Lesne, 1988). Instead, non-linear and double-linear damage rules are used for variable amplitude loads (Benkabouche et al., 2015). Irreversible deformation or residual stress builds in the tower crane structure as the variable amplitude fatigue load progresses. As a result, cracks are initiated; they propagate in the plane once they begin because of residual stresses. These irreversible local deformations and crack propagation influence the overall stability and reliability of the tower crane. To analyze the damage in the tower crane structure, total strain, residual stress, stiffness degradation, strength degradation, heat dissipation due to micro cracking, crack propagation, and the speed of sound may all be employed. In this study, we simulate the mechanical response of the crane structure by considering the stress, displacement, and damage index by applying a moving load with variable amplitude. According to the existing literature (Bandara et al., 2015; Benkabouche et al., 2015; Cantrell, 2006; Chaboche and Lesne, 1988; Fatemi and Vangt, 1998; Liu et al., 2021; Pape and Neu, 2001; Pompetzki et al., 1990; Taheri et al., 2013; Teng et al.,

2020; Wu et al., 2021), the damage rate is dependent on various variables, as expressed in Eq. (4-1). These variables are used to determine the damage rate of the tower crane.

$$dD = f(\sigma_M, \bar{\sigma}, D, \Delta\sigma)dN \tag{4-1}$$

where dD is the damage rate, σ_M is the maximum stress, $\bar{\sigma}$ is the mean stress, D is the current damage state, $\Delta\sigma$ is the stress spectrum in one cycle, and N is the number of cycles.

4.2.2.2 Finite-element analysis (FEA), degradation index, and degradation rate

Fatigue damage accumulation analysis is performed to determine the degradation rate of a crane structure for 70 years in operation by following the steps shown in Figure 4.2. The tower crane's estimated life is 50-60 years (Wiethorn et al., 2016; Wu et al., 2021); however, we chose 70 years to enhance the quality and accuracy of the DTD model. First, a FEM model specifying the geometric specifications and the material properties is developed to analyze the structural performance of the tower crane for 70 years of working conditions. Then, by considering the damage accumulation behavior of the material, moving loads with variable amplitude are applied to the tower crane cyclically to simulate the mechanical response of the crane (e.g., stress, displacement, and damage index). Figure 4.3 shows an example of damage index contours for a simulated tower crane after working for 10, 20, 30, 40, 50, 60, and 70 years. Zero indicates no damage to the crane, and one suggests that the crane model is completely damaged due to fatigue. In this case, the degradation rate (dD) is determined by subtracting the damage index from 1, indicated in Eq. (4-2).

$$dD = 1 - dI \tag{4-2}$$

where dD denotes the damage rate, and dI denotes the damage index, which varies from 0 to 1.

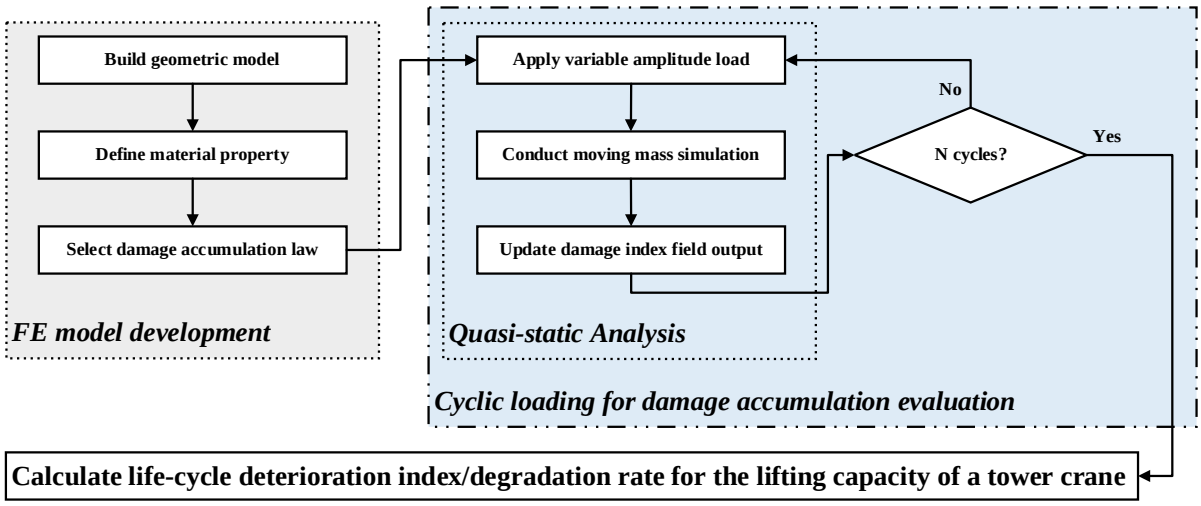


Figure 4. 2: Numerical procedure for a tower crane's fatigue damage accumulation modeling.

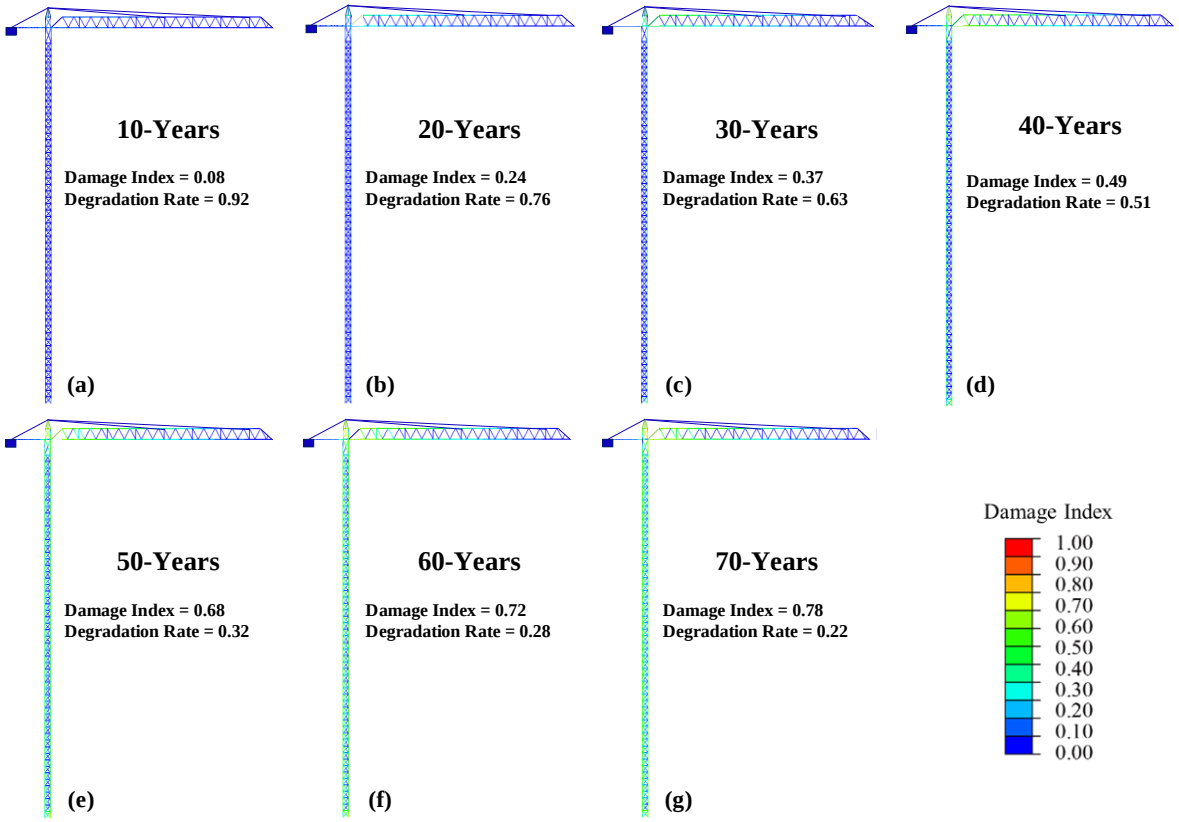


Figure 4. 3: Fatigue damage accumulation analysis: (a) 10-Years, (b) 20-Years, (c) 30-Years, (d) 40-Years, (e) 50-Years, (f) 60-Years, and (g) 70-Years.

4.2.3 A digital twin-driven model of a tower crane

This study aims to construct a DTD model for predicting degradation in LC, making intelligent decisions, and providing feedback for the lifting operation/process based on real-time data, as

illustrated in Figure 4.4. A DTD model has five parts, as indicated in Section 4.2.1. These components are fused via DT data. Multi-source, heterogeneous, and real-time data are collected from a scaled-down prototype and stored in the Firebase Real-time Database. This data is used for prediction and decision-making in the control system, including pre-processing and functional modules. The results are fed back to the physical tower crane to improve the performance and monitoring system. The control system analyzes and visualizes data for operators, decision-makers, and management teams. The proposed DTD model realizes the interaction and fusion of DTD components for monitoring degradation in LC.

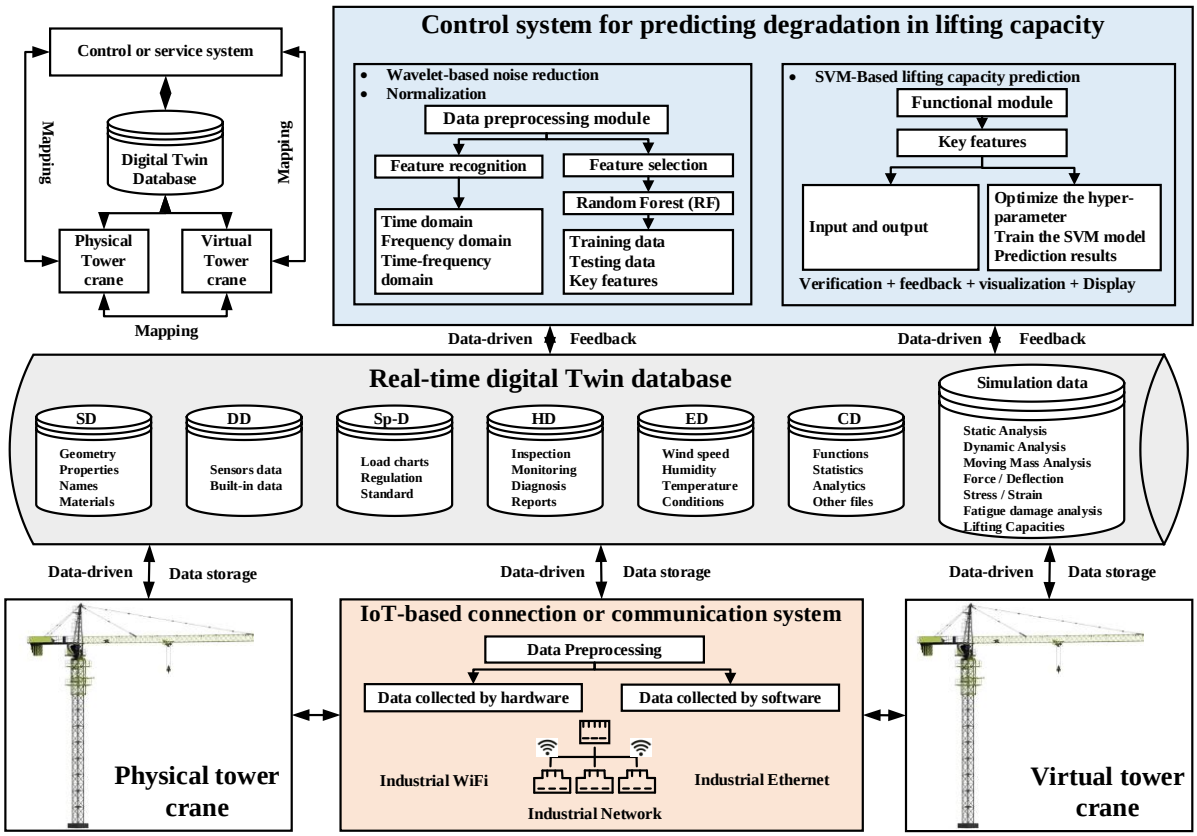


Figure 4. 4: A DTD model for predicting degradation in LC of a tower crane.

4.2.3.1 Physical tower crane

A tower crane is a high-rise structure that hoists heavy objects from one place to another on a construction site. Residual stresses accumulate in a crane structure because of high-cycle and variable amplitude fatigue loads, and LC degrades with time in service. Jiang et al. (2022)

claimed that two research areas are used to examine the tower crane's structural mechanism, operation process, and failure mechanisms: (1) modeling and simulation and (2) scale model. The finite element method (FEM) is advised as the most effective technique for analyzing the crane's internal mechanism, dynamics, and structural behavior (Wu et al., 2021). However, modeling and simulation focus on critical aspects, design optimization, and safety risk performance. Furthermore, the scale model is another structural analysis and failure mechanism research area. In contrast to the numerical model, the physical scale model conducts experiments to test the operation and control process. Therefore, it can easily detect errors and deviations in various working conditions. He et al. (2021) suggested that integrated modeling, simulation, and scale models can improve the structural analysis of tower cranes, predict failure modes, dynamic analysis, and operation processes. The preceding research serves as the foundation for the DTD model, which we propose using integrated modeling, simulation, and scale modeling in this study.

The scale model substitutes the large-size tower crane (STL420) structure. We construct a scale model based on the "similarity principle," as indicated in Eq. (4-3). The scale model contains the geometric attributes of all components, as shown in Table 4.1. The dimensions of all parts of the large-size tower crane are employed to build a scale model, as illustrated in Figure 4.5. In terms of materials, the large-size tower crane is made of stainless steel, but the scale model tower crane is made of plastic; however, we built a DTD model based on the scale model. Thus, plastic materials are defined during FE modeling.

$$S_g = \frac{L_S}{L_R} = \frac{W_S}{W_R} = \frac{H_S}{H_R} \tag{4-3}$$

where S_g is the geometric similarity factor, S is the scale model, R is the real model, and L, W, and H are length, width, and height, respectively.

Table 4. 1: Geometric similarity factors.

Structure	Scale size (L, W, H; cm)	Real size (L, W, H; cm)	Scale factor
Mast	(5.5, 5.5, 5.5)	(240, 240, 240)	1/43.6
Jib	L: 60	L: 6000	1/100
Counter jib	L: 22	L: 850	1/38.6
Rope	L: 122	L: 1650	1/13.5
Cross foundation	L*H: 34*19	L*H: 800*800	1/23.5*1/42.1
Load			
R (cm) 6000cm	500(gram)	1.089e+7 (gram)	1/21,772.4

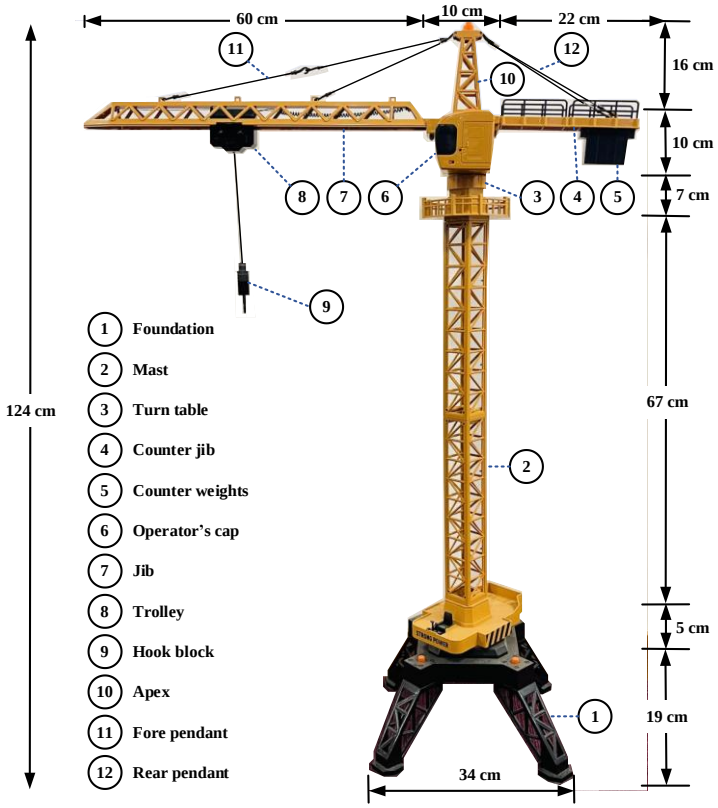


Figure 4. 5: Scale model of the tower crane and its components and dimensions.

Data is the most essential and fundamental concept in developing a DTD model. In addition to sensors, other data types could be included to enhance the quality of a DTD model, particularly in the degradation scenario. As illustrated in Figure 4.6, the tower crane generates multi-source heterogeneous data classified into six categories based on their type and nature, such as static, dynamic, specification, historical, environmental, and comprehensive data. Static data (SD) is

data that does not change with time. Although SD does not directly influence degradation, it can affect quality if altered. In contrast, dynamic data (DD) varies with time. DD provides information about the crane's dynamics, lifting operation, degradation, and real-time operation status. DD is more important than other data types for developing the DTD model. As a result, it requires high accuracy, quality, and timeliness. DD consists of sensors and built-in data. Furthermore, specification data (Sp-D) includes knowledge, standards, and load charts. Historical data (HD) refers to data from inspections, monitoring, and maintenance reports during the service life. Besides this, environmental data (ED) refer to external working conditions, such as temperature, wind, and humidity. Finally, comprehensive data (CD) is additional data other than the previously mentioned data that provides essential and specific information to the DTD model for enhancing prediction and decision-making. In addition, simulation data is required for the DTD model, which we will explain in Section 4.2.3.2. To recap, we need all the data types mentioned above at various locations to develop a DTD model to predict degradation in LC.

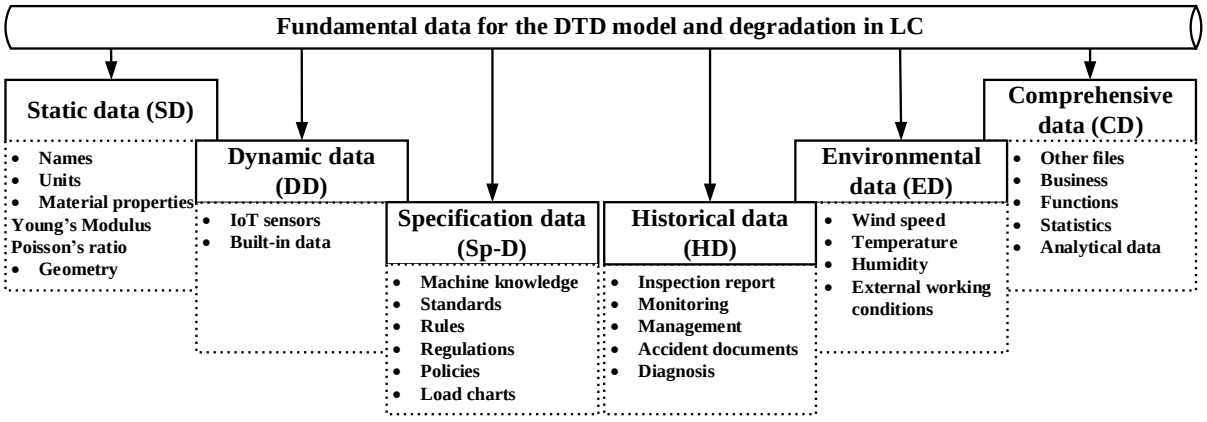


Figure 4. 6: Fundamental data for the DTD model and degradation.

4.2.3.2 Virtual tower crane

The virtual model serves as a testbed for structural and operational analysis of the tower crane. From the prospect of LC deterioration, the virtual model is a complete synchronization of the actual lifting operation of the tower crane. The ABAQUS software is used to construct the

tower crane's finite element model (FEM). A FE analysis is an effective numerical simulation for determining the capacity of structural components in the load and deflection curve. The authors considered the scale model's working conditions when analyzing static, dynamic, and moving mass on a crane model. All factors associated with lifting operation and degradation act as real-time external incentive time series data on the virtual model.

The tower crane is attached to the foundation by a baseplate and anchor bolts, which carry the bending moment and restrains four nodes and six degrees of freedom (DOF). The pin shaft connection between the mast and jib sections is considered rigid. The connection between the mast, tower cap, and pivoting support is consolidated. The jib and counter jib are attached through a hinge connection to the mast. FEA is performed using dead-load, service, wind, and concentrated loads to determine the LC of the tower crane using a force/deflection or stress/strain curve. Fatigue damage accumulation analysis assesses the damage and degradation rate for 70 years. ABAQUS software stores all lifting operation information, including the geometry of the scale components, mechanics information, operation process, and environment information. Finally, the degradation rate is integrated and manipulated with the real-time sensors data in the control system of the DTD model to predict degradation in the LC of a tower crane which will be explained in Section 4.2.3.5.

4.2.3.3 IoT-based connection and dataflow

The IoT system detects the current status of the physical tower crane. In the DTD model, the IoT system serves as a real-time data source and a safe data transfer mechanism. As depicted in Figure 4.7, the IoT system has five layers: object, sensing, communication, data management, and application. The object is the scaled-down prototype of the tower crane, which is liable for generating real-time sensors and fundamental data, as discussed in Section 4.2.3.1. The sensing layer comprises vibration/deformation/deflection and load cell sensors. The sensing layer is mainly used for data acquisition and control via a different interface, such

as OPC-UA. The communication layer accumulates and transfers all data types using serial port converters and communication modules. Furthermore, the data management layer comprises a perception, knowledge, and algorithm database. It stores and fuses essential/fundamental data translated from one form to another to assist the application layer. Finally, the application layer is the fundamental part of an IoT system. It allows communication between the OPC interface and the server. It is responsible for cleaning, associating, mining, fusing, and analyzing data across all subsystems.

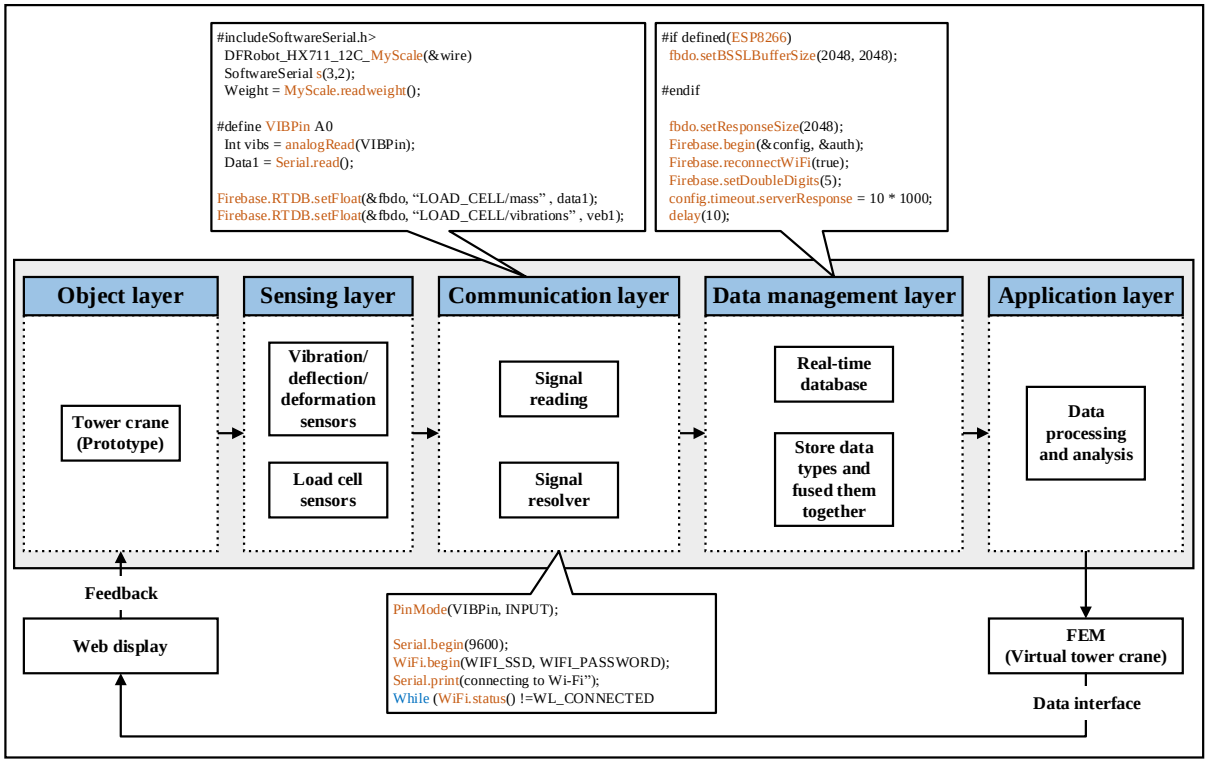


Figure 4. 7: IoT system and dataflow.

4.2.3.4 Real-time digital twin database

Advanced data conversion technology is required to integrate and manipulate sensors, fundamental, and simulation data for extracting essential features related to LC degradation. A Firebase real-time database is utilized to cover all aspects of LC deterioration. LC often relies on various factors and is described in multiple data formats with multi-source and heterogeneous characteristics.

4.2.3.5 Control system for predicting degradation

Traditionally, operators and workers are mainly responsible for monitoring and controlling lifting operations. A qualified individual follows the manufacturer's specifications and directions to ensure a safe lifting operation. The manufacturer's rated loads for a crane are based on ideal conditions. If an individual adheres to the manufacturer's loads for an extended period, the tower crane may be damaged. Therefore, a hybrid and real-time approach is required to illustrate the current status, predict the LC of a tower crane, and present the predicted LC in updated load charts and graphs. The control system is thus designed and comprises (1) pre-processing and (2) functional modules.

DT data monitoring, analysis, and evaluation are carried out in the pre-processing module. Noise is present in the signals collected from sensors. Reducing noise with the wavelet-based de-noise approach considerably impacts LC prediction. This module is also responsible for data normalization. Feature recognition should be performed using the time, frequency, and time-frequency domains to extract information from the signals. Following feature recognition, the feature selection method is applied. For further processing, the dataset is divided into training and testing datasets.

In the functional module, the important features that influence the LC of a tower crane have been selected. Based on real-time data, a DTD technique assesses the present condition of the lifting operation. However, real-time sensor data cannot predict the degradation rate in LC and the rate at which the LC decreases with time. Therefore, we performed FEA to assess fatigue damage accumulation and degradation to address this issue. Then the degradation rate is incorporated and manipulated into the database to accurately predict the LC by using an appropriate machine learning (ML) model for the RUL of a tower crane. Finally, the findings of the high precision are fed back to the physical tower crane to minimize unexpected failures

and damages while enhancing monitoring and a safe working environment in the lifting operation.

4.3 Case study

This study employs a scaled-down tower crane prototype of STL420 to evaluate the proposed DTD model's accuracy, efficacy, and applicability. There are some flaws in the prototype, but none of them interfere with our research goals. For instance, the shape of the foundation/baseplate differs from the actual crane, but the bending moment is restrained to four nodes and six DOF, much as with real cranes. In addition, the mast of the prototype has only two sections, but the mast of the actual crane has more sections, depending on height. However, this does not affect our model because joint failures are not considered in the degradation. Finally, the slewing unit and climbing frame are different in shape but identical in function.

4.3.1 Physical hardware setup

The prototype is powered by two direct-current (DC) motors that permit horizontal rotation in clockwise and counter-clockwise directions, trolley movement inward and outward, and loading/unloading through a remote control. In addition, as depicted in Figure 4.8, two sensors (load cell and piezoelectric) are attached to the prototype to measure the load and deflection/deformation. The piezoelectric sensor is attached to the end of the jib where the deflection is higher, while the load cell is mounted under the hook. The load cell is connected to the HX711 ADC, Arduino Uno, and NodeMCU ESP8266 Wi-Fi module for data collection and storage. This configuration enabled physical-virtual cranes to communicate in both directions.

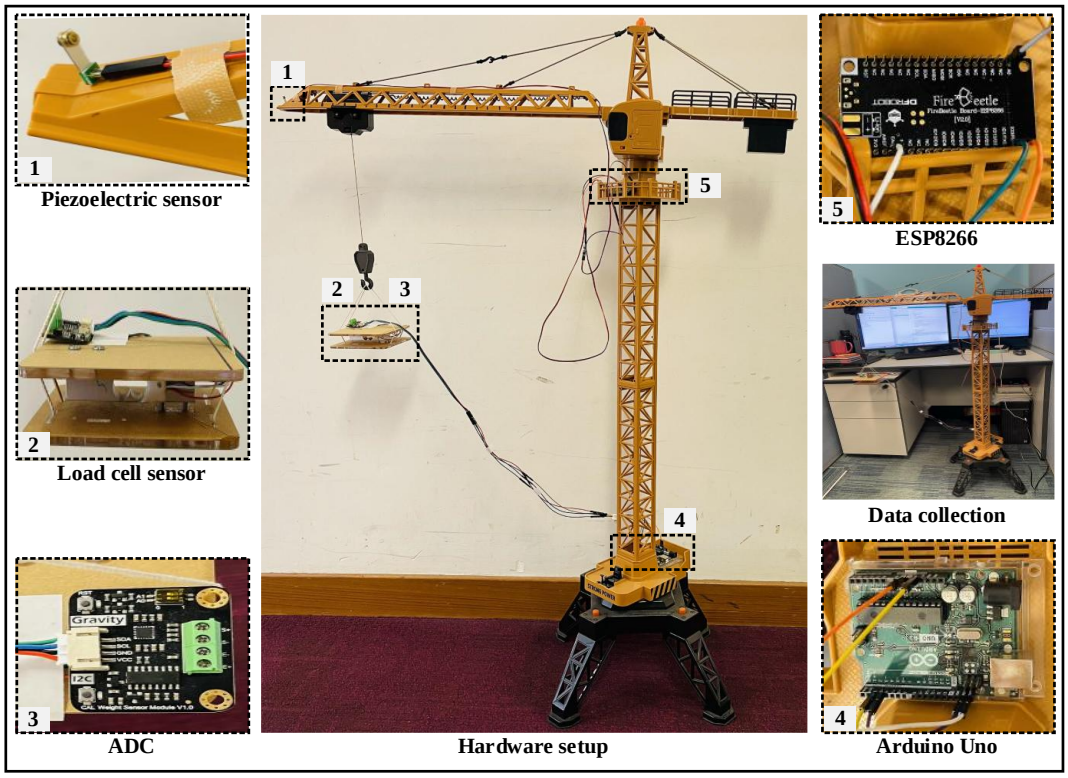


Figure 4. 8: Physical hardware setup.

4.3.2 DTD model experimental platform

The primary goal of the experimental platform is to perform lifting operations. Therefore, a scale model is assembled, and lifting operations are performed to generate and collect data. A load cell and piezoelectric sensors are installed to measure loads and deflection. Besides this, three load charts for counterweights 12, 18, and 24 tons were provided for the actual crane, and four jibs of lengths 30m, 40m, 50m, and 60m were attached to the mast. Based on the similarity principle, we employed a scale model of a 60-meter jib and 24 tons of counterweights. A load chart of the actual crane is shown in Table 4.2. Based on the scale factor, the load chart of a scaled-down tower crane is shown in Table 4.3.

As depicted in Figure 4.9, the scale model has a jib length of 60 cm, and a trolley can only move between 15cm and 50cm. Furthermore, based on scaled-down factors, we compute a load for various lifting radii and marked radius on a jib with a black marker, as shown in Figure 4.8. The calculated loads are then applied to the marked lifting radius. For each point load, three

operations are performed to measure the load and deflection: (1) loading/unloading, (2) clockwise and counter-clockwise rotation, and (3) rotation with loading/unloading. The experiment duration was 10 minutes for each point load, and we covered all three mentioned operations within 10 minutes. We conducted three tests for each point load and then averaged the results. A dataset was generated by conducting 54 experiments on 18 different locations. At the same time, the hoisting process was carried out in the virtual environment. The calculated load values from Table 4.3 were used to determine the stress/strain or load/deflection relationship. As we explained in Section 4.2.2, residual stresses accumulate, and degradation occurs if the load is applied in the cyclic mode for several years. That is why we calculated the degradation rate. Finally, sensor, simulation, and degradation rates were incorporated and manipulated to predict LC degradation.

Table 4. 2: A load chart of the actual tower crane.

Jib (m)	LC Max (t)	Lifting radius (m)														
		20	22	24	26	28	30	32	34	36	38	40	45	50	55	60
60	12	12.00	12.00	12.00	11.84	10.99	10.23	9.56	8.98	8.45	7.99	7.57	6.68	5.97	5.38	4.90
50	18	18.00	18.00	17.26	15.89	14.72	13.70	12.81	12.02	11.33	10.70	10.14	8.95	8.00		
40	24	22.66	20.52	18.74	17.23	15.94	14.73	13.75	12.89	12.12	11.43	10.85				
30	24	23.29	21.09	19.25	17.70	16.36	15.25									

Table 4. 3: A load chart of the scale-down tower crane prototype.

Jib (cm)	LC Max (g)	Lifting radius (cm)														
		20	22	24	26	28	30	32	34	36	38	40	45	50	55	60
60	500	500.0	500.0	500.0	493.3	457.9	426.2	398.3	374.1	352.0	332.9	315.4	278.3	248.7	224.1	204.1

50	750	750	750	719	662	613	570	533	500	472	445	422	372	333
		.0	.0	.1	.0	.3	.8	.7	.8	.0	.8	.5	.9	.3
40	100	944	855	780	717	664	613	572	537	505	476	452		
	0	.1	.0	.8	.9	.1	.7	.9	.0	.0	.2	.0		
30	100	970	878	802	737	681	635							
	0	.4	.7	.0	.5	.6	.4							

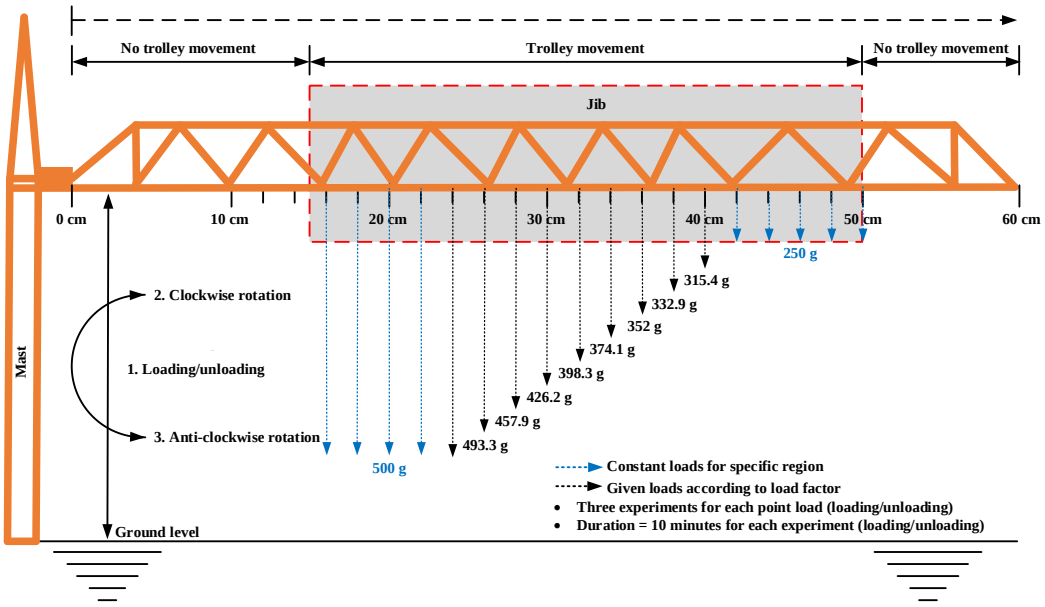


Figure 4. 9: Experimental platform for the lifting operations.

4.3.3 Data collection

As shown in Figure 4.10, the load cell and piezoelectric sensors are connected to Arduino Uno analog input pins for data collection and pre-processing. First, the raw data is measured as a voltage between 0 to 5 volts at 250 hertz. The raw data is converted into integer values ranging from 0 to 1023 using a 10-bit HX711 ADC. Second, load and deformation/deflection data are smoothed by taking an average of 20 measurements for three different weights (100g, 200g, and 300g) to eliminate noise and calibrate data. Finally, the smoothed data is multiplied by a pre-determined conversion factor to calculate load (gram) and deflection/deformation (mm) measurements.

In addition, the data collection module incorporates SD, Sp-D, HD, ED, and CD at various stages of the DTD model. For instance, geometric and material properties were considered while modeling a virtual tower crane. In the case of HD, a diagnosis report was examined for the precision and accuracy of deterioration because repairing and replacing tools affect degradation. Furthermore, during finite-element and fatigue damage accumulation analyses, wind load, temperature, humidity, and other environmental conditions were considered.

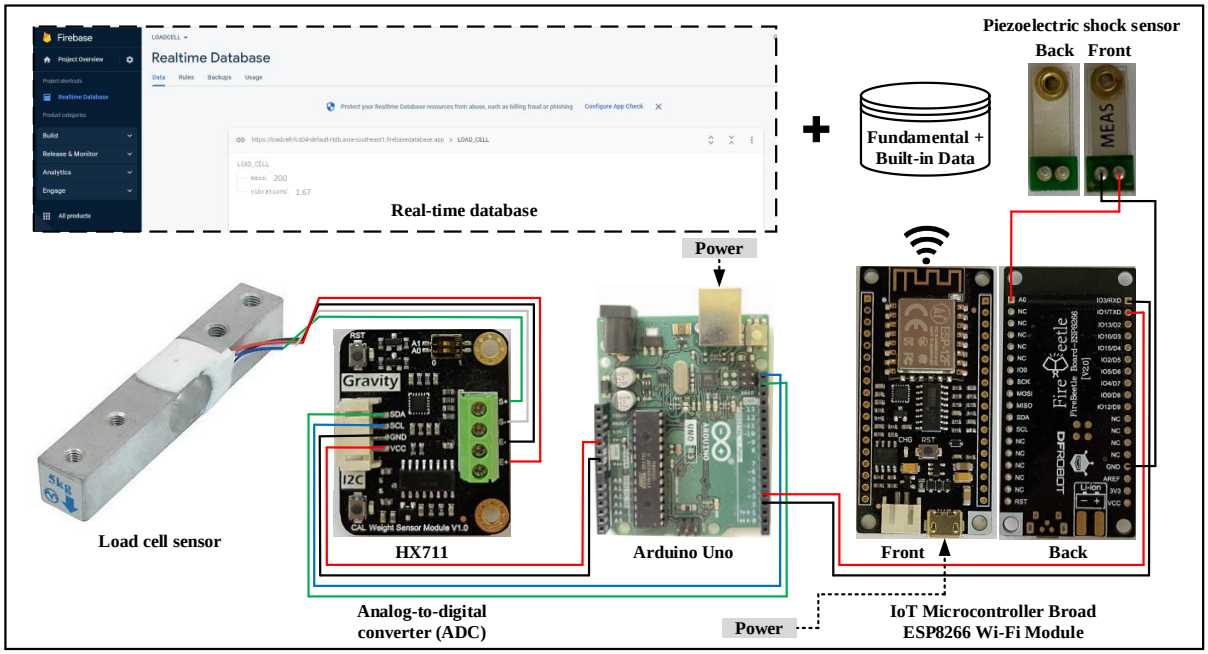


Figure 4. 10: A circuit diagram or IoT system for data collection.

4.3.4 Data transmission to the cloud

The Arduino sends the loads and vibration/deflection/deformation data to an ESP8266 via a serial port. Serial transmission takes 37 milliseconds, according to measurements. Average measurements are programmed to be delivered every second to conserve transmission bandwidth. This guarantees that data is reliably transmitted between the sensors and the cloud gateway. Furthermore, whenever an average measurement is made, the cloud gateway converts the values into JSON format and adds the timestamp.

4.3.5 FE analysis or simulation environment

The goal of simulation in a DTD model is to replicate the same process or operation in a virtual environment. ABAQUS software is employed to construct a finite-element model (FEM). As depicted in Figure 4.11, a non-linear static analysis is performed on a crane structure due to material and geometric nonlinearity. Additionally, there are variations in the boundary conditions, particularly for the boom structure. The load combination consists of a dead load, a wind load, and a lifting load. The equivalent static load is calculated by multiplying the static load by a dynamic load coefficient. In the virtual environment, we considered the yield and ultimate strength of the crane material. The data show a similar pattern, albeit with differing magnitudes. In other words, the crane structure vibrates/deflects if a cyclic load is applied. If the load and vibration persist for several years, the capacity will decrease due to the accumulation of residual stresses. We attempted to quantify and incorporate the degradation rate with the real-time data to lower the magnitude and align the real-time sensor data and simulation data.

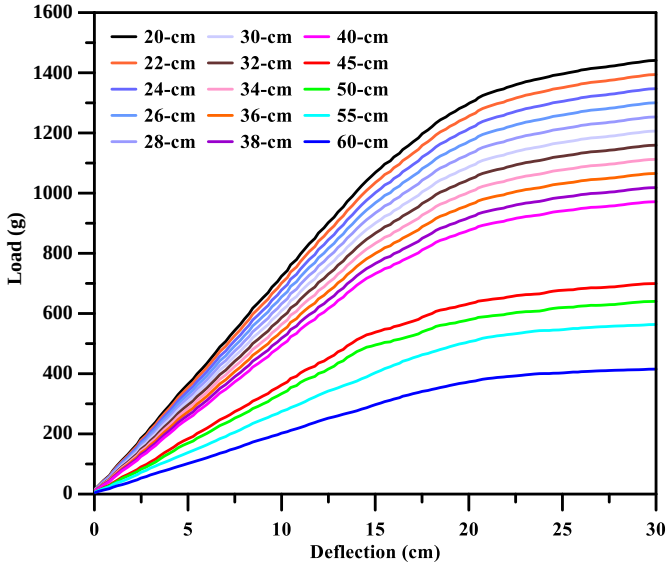


Figure 4. 11: Non-linear static analysis of the tower crane.

4.3.6 Control system and synchronization

The cloud gateway and the control module, which are connected to the cloud database, are managed by ESP8266 devices. Real-time data synchronization includes the dashboard and control module. This is accomplished using WebSockets. It allows physical and virtual cranes to communicate in both directions. The Firebase Real-time Database is employed in this study to store sensor data and control codes. The cloud gateway sends structured data to a real-time database every second. Each simulated value is maintained in the historical database to facilitate backward testing and analysis. JSON objects include six layers of nodes, which is helpful when training machine learning modules because sensor data vary in feature between operational nodes. Therefore, a classification engine based on machine learning can better use this historical data.

4.3.6.1 Historical data hosting and computation engine

Although it is challenging to construct a direct relationship between load, deflection, degradation rate, and LC, a machine learning-based classification engine can efficiently use historical data. Training and testing depend on information gathered during the idle state; however, the operation should only commence if the idle state is safe. First, the sensory data is pre-processed by grouping every 10-second interval of values. Then, for each sensor data, a rolling standard deviation is calculated to generate new features for machine learning. The most recent timestamp is the only timestamp used for the newly aggregated data. The state with the most significant number of counts is the congregated state. Delta time is overlooked because it does not contribute to the learning engine. A one-class categorization based on similarity is implemented. The historical data is then fed into a support vector machine (SVM) model. The model deduces the characteristics of usual scenarios and predicts whether the incoming input deviates from the specific occurrences. Finally, we validate the dataset by analyzing the model's

accuracy using k-fold cross-validation. To make things easier, a control system is divided into two modules: pre-processing and functional.

4.3.6.2 Preprocessing module

The pre-processing module monitors, analyses, and evaluates real-time data. The raw or sensory data contains information about the lifting operation. To extract the important features from the signal that influences the LC, feature recognition techniques like time domain, frequency domain, and time-frequency domain are employed. The key features for the LC are primarily determined, such as Load (g), deflection (mm), boom length (cm), lifting radius (cm), counterweight (g), and degradation rate. According to Roysson et al. (n.d.), degradation and LC depend on various factors, including all parameters is difficult, but optimal parameters will enhance prediction accuracy. The random forest (RF) selects the features after feature recognition. RF is the bagging model that combines the CART tree and random space. A CART tree comprises three parts: decision nodes, leaf nodes, and root nodes. The independent random sample approach is used to construct each tree. The bagging or bootstrap technique is used to sample and create the subset, and a random subset of the feature attributes is chosen based on a specific criterion. In general, RF outperforms single decision trees because the outcome of RF is selected by voting on the results of many decision trees. In this paper, we choose the essential features by employing n-estimator hyperparameters of 100, 250, 500, 750, and 1000. The dataset is divided into 70% for training and 30% for testing. The Load (g), deflection (mm), boom length (cm), lifting radius (cm), counterweight (g), and degradation rate are all critical, with values of 0.37, 0.34, 0.13, 0.14, 0.0, and 0.2, respectively. The importance of counterweight is zero here because it remains constant during lifting. Table 4.4 shows the errors and R^2 values for the RF in feature selection.

Table 4. 4: Errors and R2 values for the random forest (RF).

Model	Hyperparameter	Trees	MAE	RMSE	MAPE	R^2

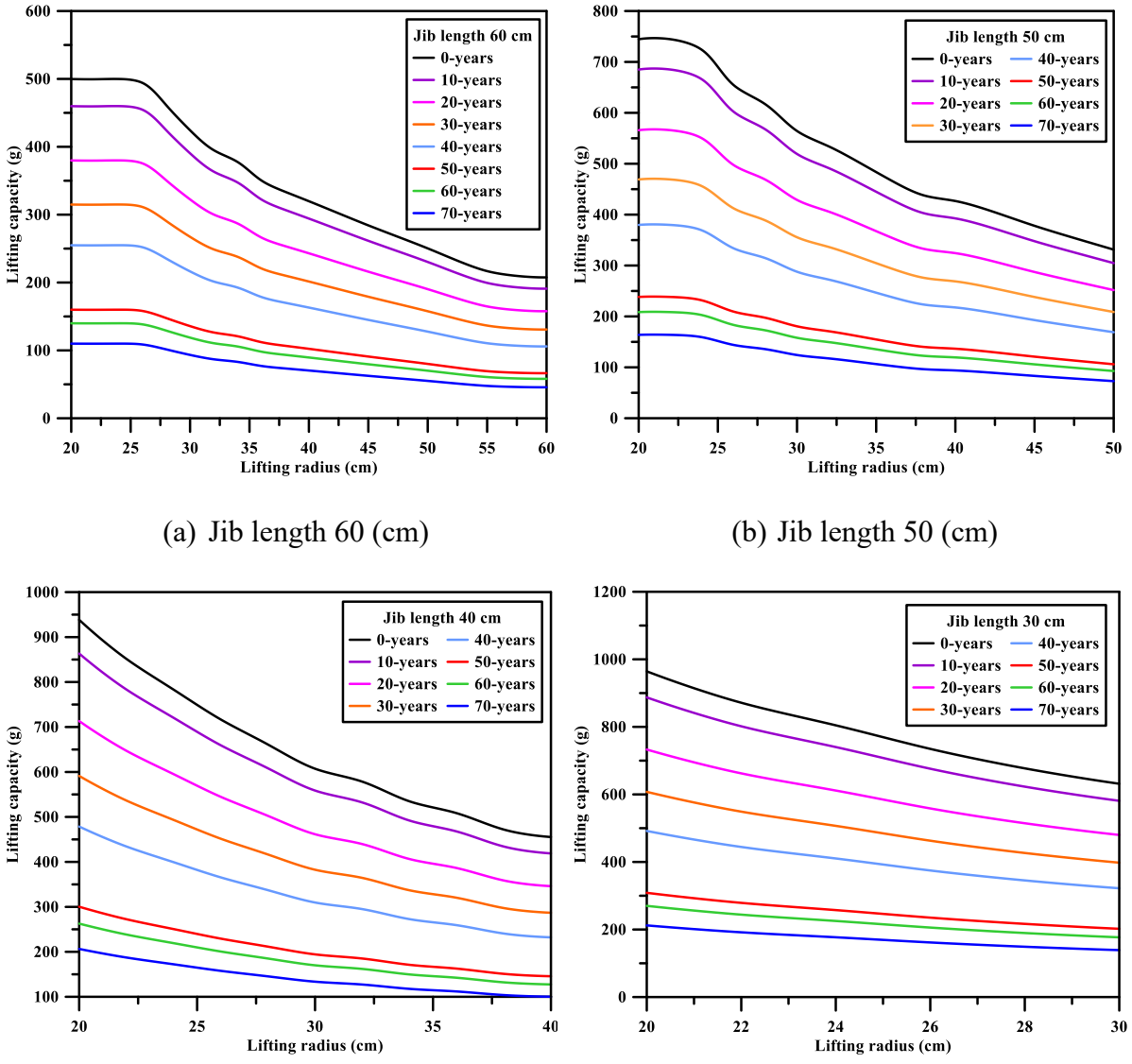
		100	0.174	0.232	2.1%	0.997
		250	0.168	0.227	2.1%	0.997
RF	n-estimator	500	0.166	0.225	2.0%	0.997
		750	0.169	0.226	2.1%	0.997
		1000	0.168	0.225	2.1%	0.977

4.3.6.3 Functional module

The LC and its degradation are complex, multi-variable, and non-linear problems; hence, using a traditional mathematical model to predict the degradation is challenging. Machine learning (ML) offers a significant advantage in terms of prediction. The SVM model predicts the degradation in LC because it can work well with high-dimensional data with a clear margin and scatter data points. To ensure the model's accuracy, the authors used 70% dataset for training and 30% for testing. The input and output variables are Load (g), deflection (mm), boom length (cm), lifting radius (cm), counterweight (g), degradation rate, and lifting capacity (g), respectively. The model's evaluation criteria are the mean-square error (MSE) and coefficient of determination (R^2). When MSE is smaller and R^2 is closer to 1, the smaller the prediction error, the better the prediction model's performance. The MSE and R^2 values are 0.2253 and 0.9973, respectively.

The evaluation criteria show that a DTD model has a certain degree of flexibility and accuracy in predicting degradation in a complex machine like a crane. The degradation rate is incorporated into the control system calculated from the fatigue damage accumulation and FE analysis. The predicted degradation in LC of a tower crane is depicted in Figure 4.12. The predicted results of LC (g) are plotted against the lifting radius (cm) for a boom length of 60cm, but it is also calculated for the 30cm, 40cm, and 50cm by multiplying the predicted results by a scale factor.

Similarly, the predicted load charts for a prototype for 0 years and 70 years of operation are shown in Table 4.5. The findings of the high precision are fed back to the physical tower cranes. Furthermore, when a scaled factor of a prototype is multiplied by the load chart, it should be equal to the actual tower crane, as shown in Table 4.6. When a degradation rate is calculated, and the load chart is updated with time throughout the tower crane's service life, the operators, management, or decision-making team can take direct action to consider more accurate LC for the crane operations or maintenance planning.



(a) Jib length 60 (cm)

(b) Jib length 50 (cm)

(c) Jib length 40 (cm)

(d) Jib length 30 (cm)

Figure 4. 12: Predicted LC of the tower crane for 70 years; (a) Jib length 60 cm, (b) Jib length 50 cm, (c) Jib length 40 cm, and (d) Jib length 30 cm.

Table 4. 5: Predicted load charts (gram) for the prototype for 0 and 70 years.

Jib (cm)	Lifting radius (cm)														
	20	22	24	26	28	30	32	34	36	38	40	45	50	55	60
60	499.8	499.5	499.9	494.1	461.2	424.3	394.9	377.6	350.4	333.7	319.8	284.2	250.3	216.8	207.6
50	744.6	744.6	723.5	654.3	616.7	564.2	534.1	501.1	466.2	437.8	426.7	378.0	331.2		
40	938.5	851.6	783.2	716.8	661.1	607.2	578.5	534.7	508.4	471.9	455.3				
30	964.3	871.4	804.5	734.8	677.1	631.5									
Jib (cm)	Lifting radius (cm)														
	20	22	24	26	28	30	32	34	36	38	40	45	50	55	60
60	110.0	109.9	110.0	108.7	101.5	93.3	86.9	83.1	77.1	73.4	70.4	62.5	55.1	47.7	45.7
50	163.8	163.8	159.2	144.0	135.7	124.1	117.5	110.3	102.6	96.3	93.9	83.2	72.9		
40	206.5	187.4	172.3	157.7	145.4	133.6	127.3	117.6	111.9	103.8	100.2				
30	212.2	191.7	177.0	161.7	149.0	138.9									

Table 4. 6: Predicted load charts (ton) for the real tower crane based on a scale factor for 0 and 70 years.

For 0-years:

Jib (m)	Lifting radius (m)														
	20	22	24	26	28	30	32	34	36	38	40	45	50	55	60
60	11.9	11.9	11.9	11.8	11.0	10.1	9.4	9.0	8.4	8.0	7.6	6.8	6.0	5.2	4.9
50	17.8	17.8	17.3	15.7	14.8	13.5	12.8	12.0	11.1	10.5	10.2	9.0	7.9		
40	22.5	20.4	18.7	17.2	15.8	14.5	13.8	12.8	12.2	11.3	10.9				
30	23.1	20.9	19.3	17.6	16.2	15.1									

For 70-years:

Jib (m)	Lifting radius (m)														
	20	22	24	26	28	30	32	34	36	38	40	45	50	55	60
60	2.6	2.6	2.6	2.6	2.4	2.2	2.1	2.0	1.9	1.8	1.7	1.5	1.3	1.1	1.1
50	3.9	3.9	3.8	3.5	3.3	3.0	2.8	2.6	2.5	2.3	2.3	2.0	1.7		

40	5.0	4.5	4.1	3.8	3.5	3.2	3.1	2.8	2.7	2.5	2.4
30	5.1	4.6	4.2	3.9	3.6	3.3					

4.4 Validation

A DTD model is conducted with sample data as a constraint. To execute the unpredictability and chance of the experiment, the degradation in LC of a tower crane is evaluated using a k-fold cross-validation approach to validate the precision of a DTD model. The sample data are trained and predicted using k-5 cross-validation and SVM. The accuracy of predictions is 0.97(R^2), nearly the same as the typical SVM model. As a result, the proposed DTD model is convenient, efficient, and based on real-time data.

4.5 Discussion

This section discusses the DTD model results, including physical tower crane, virtual tower crane, IoT-based connection, data storage, and control system. Firstly, we used a scale-down tower crane prototype, fitted load cell, and piezoelectric sensors to collect load and deflection data. The prototype also generates fundamental data, which we consider at various stages of a DTD model. Existing approaches, such as empirical approaches (Chaboche and Lesne, 1988), non-destructive technologies (Przybyłek and Więckowski, 2022), and numerical simulations (Bandara et al., 2015; Cantrell, 2006; Fatemi and Vangt, 1998), necessitated the use of idle and unreal models to conduct experiments to determine failure mechanisms. Whereas data-driven models (Lee et al., 2019; Tran et al., 2012) are data-intensive and need data for at least one year. Based on the similarity principle, the scaled model substitutes the actual tower crane for assessing structural analysis, failure mechanism, dynamics, operation, and control process as realistically as possible. For instance, Jiang et al. (2022) used a prototype of a tower crane to study stability analysis and hoisting safety. Similarly, He et al. (2021) used a prototype to study the effect of wind on cranes. The proposed DTD model offers various benefits over the existing

methods because, in the DTD model, we integrated a scaled model, sensors, and simulation data to predict LC degradation.

Secondly, in a DTD model, the objective of the simulation is to replicate the lifting operation in a virtual environment. Therefore, the simulation of the virtual body may be adjusted, edited, updated, and calibrated when the physical and virtual bodies match in LC. In our study, the load and deflection are the same in the physical and virtual cranes. Still, there is a minor discrepancy between the capacity of materials in physical and virtual bodies because of safety issues on a construction site and experimental platform, as illustrated in Figure 4.13. ABAQUS uses non-linear static analysis on a scaled model to calculate material capacity as load/deflection. Unlike load charts, manufacturing companies provide LC for specific cranes based on material properties. Nonetheless, we have some safety margin in the construction or experimental platform to avoid unexpected failures. The collisions between physical and virtual LC are caused by a factor of safety (FOS). In Figure 4.13, the area under the blue line is considered a safe/working zone. The zone will be considered as dangerous if it surpasses the blue line. The area under the red line is considered a damage zone, while it is a failure zone beyond the red line, indicated in white. The physical and virtual data were in a similar pattern, albeit with differing magnitudes. We quantify the degradation rate and incorporate it with the real-time data to lower the magnitude and align the real-time sensor and simulation data (Benkabouche et al., 2015; Dong et al., 2021a, 2021b; Fatemi and Vangt, 1998; Taheri et al., 2013; Wu et al., 2021). Yang et al. (2022) used a similar approach; they multiplied wear data with the field-measurable precision index value. To summarize, we built a DTD model based on working loads and deflection; if the working loads and deflection continue for several years, residual stresses will degrade the material capacity; therefore, we calculated the residual stresses and incorporated them with the operational values to predict degradation.

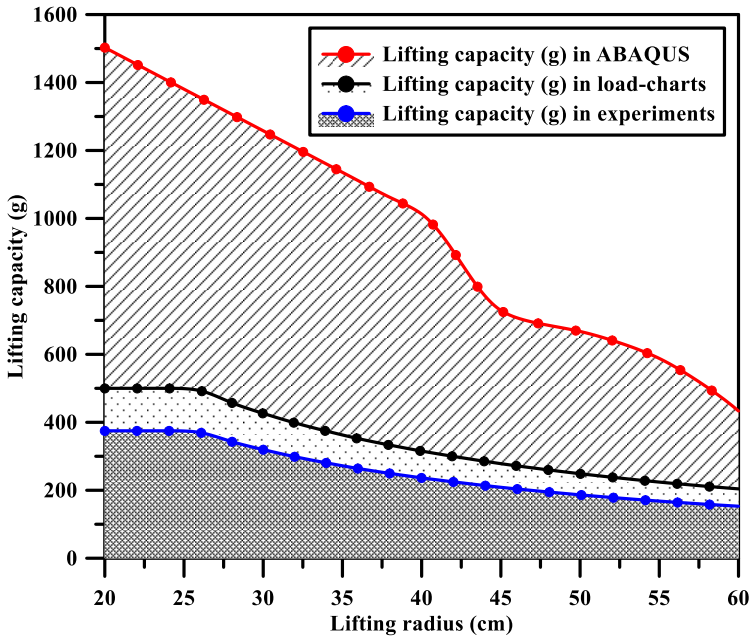


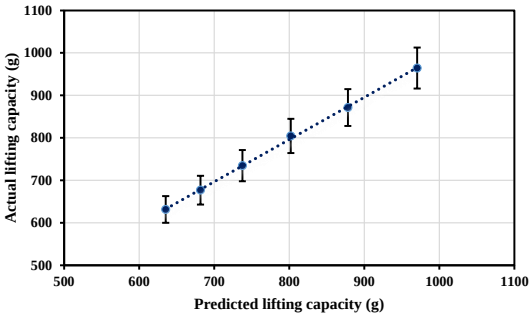
Figure 4. 13: Physical and virtual cranes load comparison and their zones.

Similarly, in a DTD model, data flow is automated and stored in a real-time database. A bidirectional connection exists between virtual and physical cranes, although digital shadow lacks this feature (Sepasgozar, 2021). In addition, LC and degradation rely on various factors and are described in multiple data formats with multi-source and heterogeneous characteristics. Therefore, advanced data conversion technology is required to integrate and manipulate sensors, simulations, and fundamental data for extracting features related to LC and degradation. Furthermore, all components of the proposed DTD model are fused and mapped to detect/resist the real-time changes in a physical space and feedback from the control system. It is also considered the secure connection and data transmission from the physical crane to the virtual crane and vice versa.

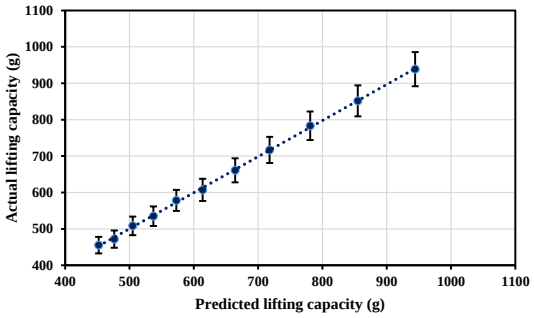
Finally, the control system predicts degradation in LC based on real-time data. It is a hybrid mode and the most fundamental part of a DTD model, which is used to analyze and visualize real-time changes, make decisions, and provide feedback for improvement. The control system is responsible for pre-processing, de-noise, calibration, normalization, and raw and real-time data evaluation. It also includes feature recognition and selection. In DT technologies, sensor

and simulation data should match and cover optimal parameters associated with a particular activity and improve prediction after integrating and adjusting these data. In this paper, we predicted degradation in LC of a tower crane for 70 years and displayed the results in graphs and load charts. Furthermore, using scale and similarity factors, we also predicted degradation for various scale-down jib lengths, including 30cm, 40cm, 50cm, and 60cm. Roysson et al. (2016) employed ANN to predict the LC of a mobile crane; however, the average deviation between the actual and predicted LC was 1.3, but in our case, the average deviation between the actual and predicted LC is 2.6, which is high than the previous research. We used sensors, simulation, degradation rate, and six input parameters; they used data from load charts and only three input parameters.

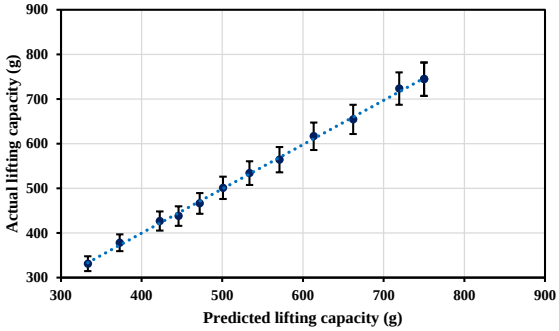
Similarly, Yang et al. (2022) used a DTD approach to predict the performance degradation in the transmission unit of the CNC machine tool with an average error of 4.33%. In this paper, we selected MSE and R^2 as the evaluation criteria for the DTD model. The MSE and R^2 of the prediction model are 0.2253 and 0.9973, respectively. We validate our model by comparing the results with the traditional SVM model, and the average accuracy of predictions is almost 99%. As depicted in Figure 4.14, the average deviation between the actual and predicted LC for the scaled-down jib lengths of 30cm, 40cm, 50cm, and 60cm are 2.9, 2.6, 3.19, and 1.8, respectively. The average deviation for all these jibs is 2.6. As shown in Figure 4.15, we multiplied the scale factor of a prototype with the projected LC for jib lengths of 30cm, 40cm, 50cm, and 60cm, and the resulting deviations are 0.04, 0.07, 0.06, and 0.07, respectively.



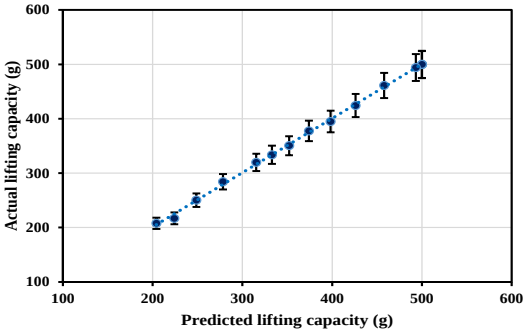
(a) Jib length 30 cm



(b) Jib length 40 cm

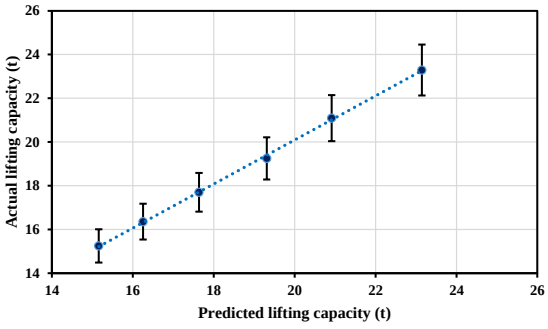


(c) Jib length 50 cm

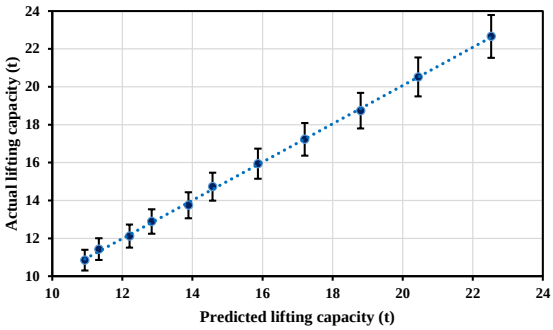


(d) Jib length 60 cm

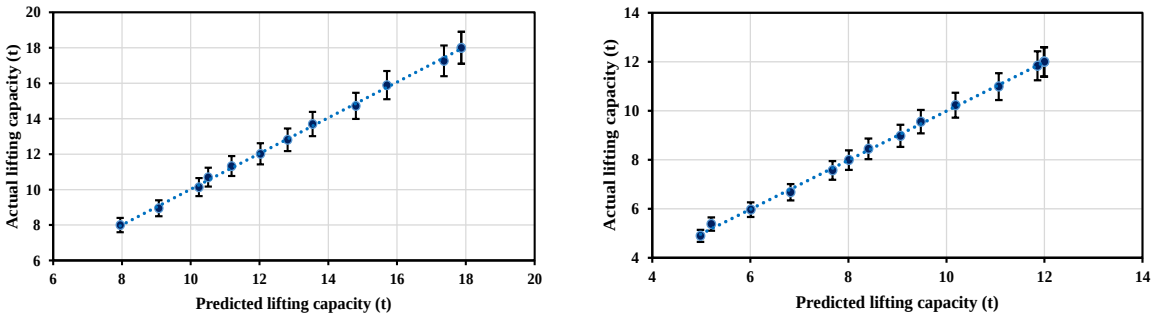
Figure 4. 14: Deviations between the actual and predicted LC for a scaled-down tower crane prototype: (1) Jib length 30cm, (b) Jib length 40cm, (c) Jib length 50cm, and (d) Jib length 60cm.



(a) Jib length 30 cm



(b) Jib length 40 cm



(c) Jib length 50 cm (d) Jib length 60 cm

Figure 4. 15: Deviations between the actual and predicted LC for a real tower crane: (1) Jib length 30cm, (b) Jib length 40cm, (c) Jib length 50cm, and (d) Jib length 60cm.

4.6 Chapter Summary

The chapter introduces a Digital Twin-Driven (DTD) framework to predict the degraded lifting capacity (LC) of aging tower cranes. The DTD approach integrates real-time data from physical cranes with numerical simulations and machine learning models to monitor and predict the effects of structural fatigue and degradation over time. This innovation addresses a critical gap in existing safety systems, which fail to account for aging-induced LC deterioration, thereby enhancing operational safety and efficiency.

A scaled-down tower crane prototype is utilized to validate the DTD framework, demonstrating high accuracy with mean squared error (MSE) and coefficient of determination (R^2) values of 0.2253 and 0.9973, respectively. Through fatigue damage accumulation modeling and finite element analysis (FEA), the study generates predictive load charts for crane operation over decades, enabling safer load application and preventive maintenance planning.

The framework comprises a physical-virtual synchronization via IoT-enabled real-time data transmission, control systems for decision-making, and a feedback mechanism to improve crane performance. Using machine learning models such as Random Forest (RF) for feature selection and Support Vector Machine (SVM) for prediction, the DTD system offers robust degradation monitoring and prediction capabilities.

This research provides a scalable solution for real-time safety monitoring of aging tower cranes, reducing risks of unexpected failures and enhancing decision-making in construction management. Future work aims to incorporate more sensors, automate system integration, and refine the physical-virtual fidelity for further improvements. The findings underscore the transformative potential of digital twin technologies in predictive maintenance and safety optimization for the construction industry.

CHAPTER 05

AUTOMATED SAFETY RISK ASSESSMENT FOR CRANE OPERATIONS USING CASCADE LEARNING

5.1 Introduction

Construction sites vary in size, layout, and working conditions; however, they generally include workers, equipment, and materials (Shen et al., 2016; Teizer and Cheng, 2015). Heavy machinery and tools are essential for critical and complex activities to meet schedules (Zhou et al., 2018), but they pose risks if failures occur (Duarte et al., 2021; Shin et al., 2024). Accidents often stem from unsafe operation, improper use, lack of awareness, education, or skills (Gürcanli et al., 2015; Jung et al., 2022; Lee et al., 2020). Comprehensive monitoring and assessment of machine operations are crucial for understanding the causes of accidents and preventing them (Xiao and Kang, 2021; Zhang and Ge, 2022). Despite safety regulations and training, accidents continue to occur, leading to severe injuries, fatalities, and financial losses. Current safety measures often rely on manual judgments, which can be error-prone, delayed, or inconsistent (Hinze et al., 2016; Jeelani et al., 2016; Rahim Abdul Hamid et al., 2019; Shafique and Rafiq, 2019). Therefore, automated, advanced, accurate, reliable, and adaptive safety procedures are needed to identify, assess, evaluate, and monitor operational activities in real-time.

The Bureau of Labor Statistics reported that the United States recorded 5,486 fatal work injuries in 2022, reflecting a 5.70% increment from 2021. The fatal work injury rate rose to 3.70 per 100,000 full-time equivalent (FTE) workers, with worker fatality occurring roughly every 96 minutes, compared to every 101 minutes in 2021 (Bureau of Labor Statistics, 2023). Among these fatalities, 170 were attributed to being struck-by machines and 59 were due to incidents involving being caught-in, trapped-in, compressed-by, or crushed-by machines (Brown et al., 2021). Additionally, 79,660 nonfatal work-related injuries have been reported,

with a significant portion resulting from contact with machines, falls, slips, trips, and overexertion (Brown et al., 2020). Machine-related struck-by incidents remain a leading cause of construction injuries globally (Hinze et al., 2005), with high fatality rates in Hong Kong (19.98%), Japan (6.54%), Singapore (5.40%), and Australia (3.20%) (Chiang et al., 2017). Hong Kong reported 3,065 construction accidents in 2022, with slips, falls, and incidents involving machinery as significant contributors (ISSH, 2024). Approximately 91.80% of construction workers operate near heavy machinery, making them vulnerable to hazards such as being caught-in, trapped-in, compressed-by, crushed-by, and struck-by equipment (Betit et al., 2022). Among these incidents, there were 47 crane accidents in 2022, up from 39 in 2021. Of these, 38.00% occurred during operation, 31.00% during climbing, assembling, or dismantling, and 23.00% were due to other factors. The critical importance of crane safety risk assessment during operation underscores the need for Smart Site Safety Systems (4S) to review and enhance current safety measures in the construction industry (Gu et al., 2022; Niu et al., 2016).

4S represents a construction site safety management proposal integrating computer vision and deep learning technologies (Awolusi et al., 2018; Johnson, 2002; Mustafa et al., 2023). This innovative system operates through three key elements: smart monitoring devices that capture and identify potential risks, networked data transmission systems, and centralized platforms that process visual information to coordinate responsive actions (Fang et al., 2020; Paneru and Jeelani, 2021; Seo et al., 2015; Xu et al., 2021). While traditional safety protocols rely heavily on pre-construction risk assessments, 4S enables real-time hazard management during the active construction phases. This technological advancement significantly shifts from conventional practices by providing immediate detection and response capabilities for operational activities (Fung et al., 2012). The system leverages advanced computer vision algorithms to perform multiple critical functions simultaneously: detecting potential hazards,

tracking movement patterns, recognizing unsafe behaviors, and classifying risk levels in real-time (Hung and Su, 2021; Khallaf and Khallaf, 2021). This comprehensive approach is particularly vital, given the dynamic and complex nature of modern construction environments (Sabuhi et al., 2021; Singh et al., 2024; Tylman et al., 2010).

As construction projects become increasingly sophisticated and more extensive in scale, the deployment of heavy machinery has grown proportionally to meet the productivity demands (Hung and Su, 2021a; Liu et al., 2021). This intensified machinery usage inherently increases operational risks. Since accidents involving heavy construction equipment typically result in more severe consequences than other construction incidents (Shao et al., 2019), implementing robust safety measures is paramount. Current research on computer vision-based safety systems primarily focuses on specific risk factors, such as collision prevention and unsafe operational practices (Fang et al., 2018). However, a truly effective safety management system must go beyond mere identification to incorporate a thorough risk assessment and informed decision-making processes. The existing research landscape shows a notable imbalance; while considerable attention is given to hazard identification and classification, there remains a significant gap in developing quantitative operational risk assessment methodologies (Fang et al., 2019; Schieg, 2006). This limitation potentially compromises the effectiveness of construction site safety measures and risk management strategies.

Recent deep learning models based on Generative Adversarial Networks (GANs), Recurrent Neural Networks (RNNs), and convolutional neural networks (CNNs) have demonstrated reliable performance in construction applications, particularly in detecting predefined entities in video footage in near real-time. While researchers have attempted to enhance or modify specific components of GANs, RNNs, and CNNs, the overall network structures remain unchanged. A significant limitation of current safety monitoring approaches lies in their reactive nature, as they primarily focus on detecting unsafe behaviors only after they become

apparent risks (Xiao et al., 2021). This retrospective approach overlooks the numerous latent risk factors at construction sites, creating potential blind spots in safety management. The practice of triggering alerts only when hazardous activities reach critical levels is fundamentally flawed from a preventive perspective (Fang et al., 2019; Xiao et al., 2021). A proactive and holistic approach to safety risk assessment is necessary, encompassing visible hazards and potential risks still in their dormant state. Such comprehensive oversight would enable construction managers to strategically identify areas requiring enhanced safety supervision and optimize resource allocation for preventive measures (Schieg, 2006). Moreover, this forward-thinking approach would strengthen crisis response capabilities, allowing safety teams to intervene before situations escalate to accidents. By identifying and addressing risk factors in their early stages, construction sites can either entirely prevent incidents or substantially reduce their impact when they do occur. For example, Fang et al. (2018) employed an Improved Faster R-CNN (IFaster R-CNN) to detect workers and equipment more accurately than other advanced detection methods. Despite these achievements, computer vision and deep learning face several challenges, including the requirement of large-annotated datasets, manual or semi-automated feature extraction, high computational demands, and issues of robustness and adaptability (Brozovsky et al., 2024; Dobrucali et al., 2024; Pandey et al., 2024). Researchers have developed specialized datasets to address some of these challenges and employed Transformer Networks for automated image-feature extraction. Notable examples include the work of Tajeen and Zhu, (2014), Lee et al. (2022), and Xiao and Kang, (2020), who developed annotated datasets for various construction machines. Gao et al. (2024) utilized an augmented weighted bidirectional feature pyramid network to detect marine objects.

Despite recent advancements, the development of computer vision models tailored to diverse construction environments remains an ongoing challenge, particularly in terms of efficiency,

reliability, accuracy, and adaptability (Al-Faris et al., 2020; Fang et al., 2023). Researchers are actively investigating the application of transformers and cascaded learning frameworks to improve the precision, robustness, and scalability of object-detection and classification systems (Martin et al., 2023). For instance, Kim et al. (2024) integrated various models into a cascaded framework to enhance the identification and classification of unsafe actions involving equipment in low-quality images, resulting in improved accuracy. Hou et al. (2023) utilized computer vision-based systems to compute safety risks and visualize construction sites. However, comprehensive safety risk assessments require more than just detection and computation systems for unsafe operations; they also require risk identification, evaluation, control, mitigation, monitoring, and documentation. Research gaps persist in this area due to the absence of automated systems for such integrations. Current computer vision models also face challenges related to high computational demands, inconsistent reliability, limited accuracy, and difficulties adapting to diverse construction settings. Furthermore, incorporating environmental conditions and varying equipment types to enhance model robustness presents a significant challenge. The labor-intensive nature of construction work and reliance on manual judgment add further complexity to integrating these advanced technologies. Therefore, this study aims to employ a cascade learning method to automate safety risk assessment in crane lifting operations.

5.2 Methodology

5.2.1 Overview of research methodology

This section presents a comprehensive methodology for automated safety risk assessment for crane operations using cascade learning. As depicted in Figure 5.1, the framework encompasses multiple integrated phases: surveillance-based data collection, image pre-processing, risk identification, assessment, evaluation, control, monitoring, documentation, and validation, all orchestrated through a cascade learning pipeline.

The methodology initiates with surveillance data collection from construction sites featuring crane operations (mobile and tower cranes). The dataset comprises CCTV footage sourced from YouTube, systematically segmented into operational activity sequences and subsequently converted into frame-based datasets. The data collection protocol emphasizes acquiring diverse crane operation scenarios, specifically targeting safe operations, unsafe operations, near-miss incidents, and incidents in dynamic construction environments. The image pre-processing phase employs a Super-Resolution Generative Adversarial Network (SRGAN) model to address the inherent limitations of surveillance footage quality. This approach, following Kim et al. (2024), enhances low-resolution frames to high-resolution outputs, which are crucial for processing blurry CCTV footage. The risk identification phase leverages state-of-the-art transformer-based models: Real-Time Detection Transformer-Large (RT-DETR-L), self-DIstillation with NO labels (DiNOv2), and Vision Transformer (ViT). These models optimize object detection and classification within dynamic surveillance contexts. The risk assessment phase transforms confidence levels derived from activity detection into probability metrics. The probability and impact of identified risks are analyzed using the risk levels defined by Hou et al. (2023), resulting in comprehensive risk values for each operational activity. The risk evaluation framework implements threshold-based risk prioritization, where probability and impact scores determine risk criticality. Activities exceeding predetermined thresholds are flagged for immediate intervention. This feeds into the control and monitoring phase, where real-time warning systems alert operators to potentially hazardous situations, enabling prompt risk mitigation actions. The methodology incorporates continuous monitoring of crane operations and iterative review of risk management protocols to ensure sustained effectiveness. All processes, findings, and interventions are systematically documented, maintaining a comprehensive audit trail of risk assessment activities. Finally, validation was conducted for the risk assessment models and risk values to evaluate the applicability of the research in real

construction projects. This integrated approach automates safety risk assessment procedures, advances 4S technologies, and has the potential to enhance safety standards.

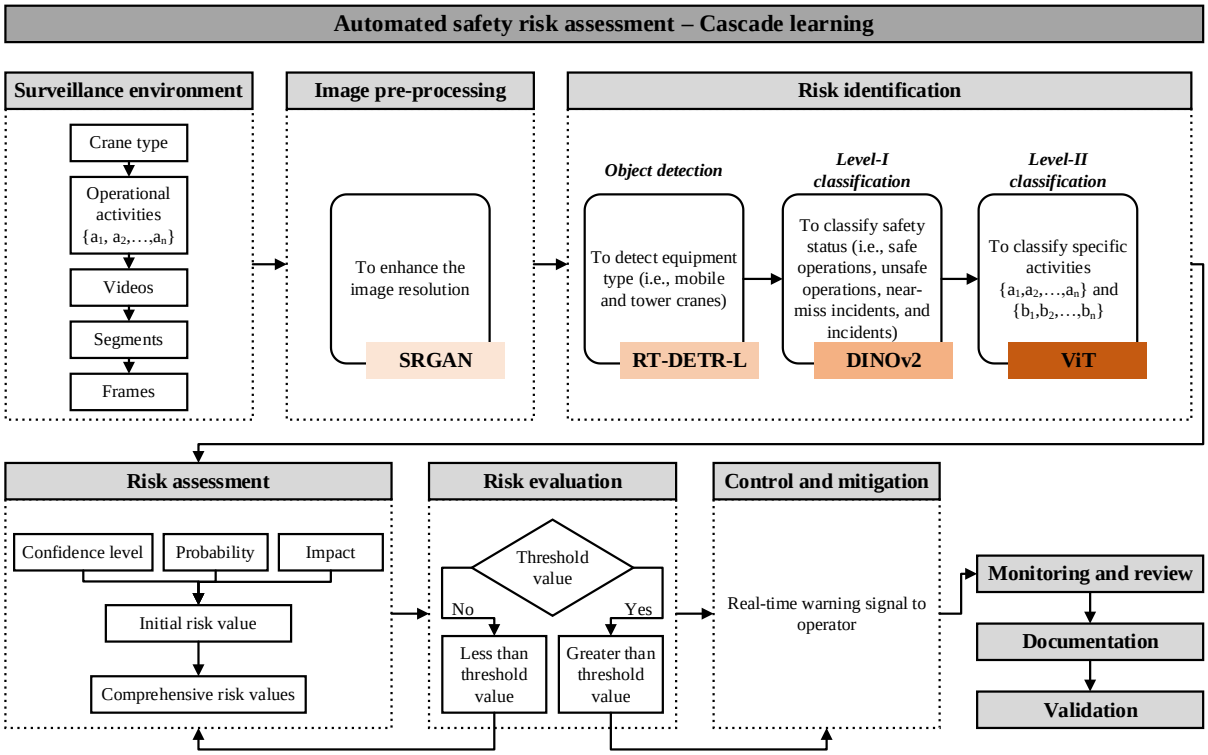


Figure 5. 1: Overview of the research methodology.

5.2.2 Data

5.2.2.1 Data collection

Our research dataset comprises a curated selection of YouTube videos depicting safe operations, unsafe operations, near-miss incidents, and incidents involving mobile and tower cranes. We selected these specific crane types due to their common use across construction sites and involvement in diverse activities. The dataset is carefully balanced to include safe operations, unsafe operations, near-miss incidents, and incidents, providing a realistic representation of real-world scenarios. Table 5.1 outlines in detail the identified activities for each crane type. These classifications were integral to our annotation process and served as a foundation for training our cascade model to enhance on-site safety. To ensure the validity of our proposed method, we developed the dataset through a systematic approach. Each video was

thoroughly evaluated to confirm that it captured the distinct actions listed in Table 5.1. Our selection process was conducted randomly and without bias, with a primary focus on the relevance of the content to our research objectives.

Table 5. 1: Overview of the selected activities in mobile and tower crane operations.

Equipment	Classes	Safety status	Subclasses	Activities	Description
Mobile crane	A	Safe operations	a_{11}	Outriggers	Four outriggers fully extended with adequate plates on flat and firm ground
			a_{12}	Movement	Movement on the road with proper posture
			a_{13}	Lifting operation	Lifting operation with proper posture
			a_{14}	Safe access to the deck	For example, a ladder to the deck of a crane
		Unsafe operations	a_{21}	Operating on unsafe terrain	Operating on soft or sloppy ground
			a_{22}	Unstable lifting operation	For instance, dragging the loads, abruptly swinging, or not correctly rigging
			a_{23}	Unclear division of work area	For instance, a lack of safety fence for the lifting/traveling zone
			a_{24}	Unsafe access to a working crane	For example, unsafe access to the cabin due to any barrier
			a_{25}	Operating near power lines	Distance should be according to safety standards
		Near-miss incidents	a_{31}	Potential collision	For example, overlapping or potential collision with other cranes
			a_{32}	Potential struck-by	For example, cross-operation
			a_{33}	Load tipping or shifting	Sudden load shifts that temporarily destabilize the crane
		Incidents	a_{41}	Collapse	Structure failures
			a_{42}	Tipping over	Crane becomes unstable and falls
			a_{43}	Fall of crane jib/boom/other parts	Fall of any components
			a_{44}	Fall of loads	Fall of loads due to rope or other reasons
			a_{45}	Collision	Collision of crane with other cranes

			a_{46}	Struck-by	Crane struck with surrounding environments except for other cranes
			a_{47}	Struck-by objects	Struck by any falling objects
Tower crane	B	Safe operations	b_{11}	Assembling/dismantling	Using safety straps such as PPE
			b_{12}	Erection	Mast erection
			b_{13}	Lifting operation	--
			b_{14}	Safe access to the deck	Proper PPE provided to the operator to access the cabin
		Unsafe operations	b_{21}	Inadequate use of PPE	During assembling/dismantling process
			b_{22}	Unstable lifting operation	--
			b_{23}	Unclear division of work area	--
			b_{24}	Unsafe access to a working crane	--
			b_{25}	Operating near power lines	--
		Near-miss incidents	b_{31}	Potential collision	--
			b_{32}	Potential struck-by	--
			b_{33}	Load tipping or shifting	--
		Incidents	b_{41}	Collapse	--
			b_{42}	Tipping over	--
			b_{43}	Fall of crane jib/boom/other parts	--
			b_{44}	Fall of loads	--
			b_{45}	Collision	--
			b_{46}	Struck-by	--
			b_{47}	Struck by objects	--

"--" indicates that the class has already been defined in the mobile crane case.

Figure 5.2 depicts a sequence of activities involving mobile cranes, which are essential for hoisting heavy modules on construction sites. The figure includes frames from videos in our dataset, capturing the crane as it performs various tasks in different environments. These images highlight the crane's versatility and illustrate the safety conditions of its operations, which we classified as safe operations, unsafe operations, near-miss incidents, and incidents in our study.

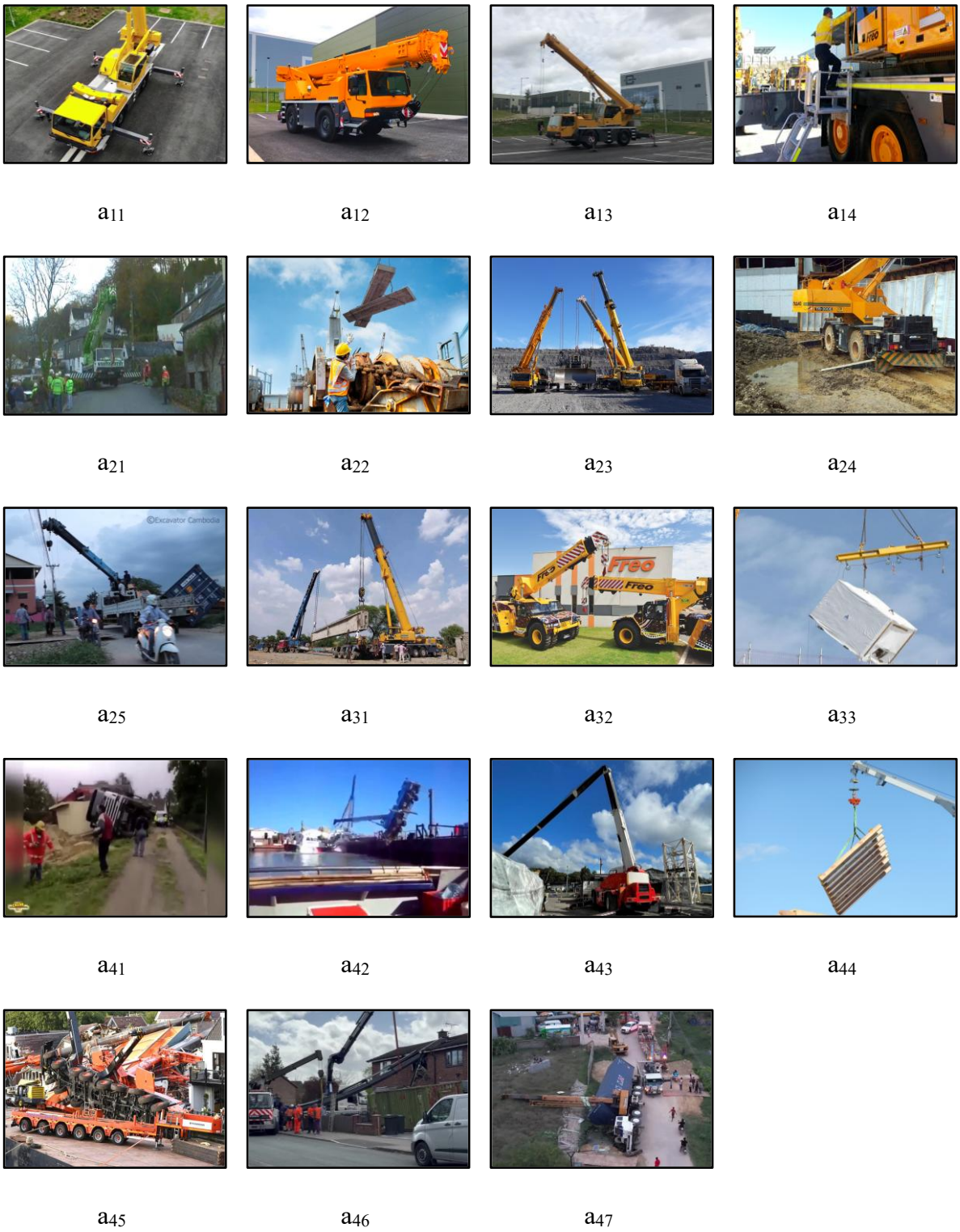


Figure 5. 2: Activity categories of mobile crane.

Figure 5.3 offers a visual compilation of tower crane activities on construction sites. The images, sourced from our video dataset, capture the crane engaged in various operations,

including safe operations, unsafe operations, near-miss incidents, and incidents. This compilation illustrates the breadth of tower crane operations and their associated safety risks.

b_{11} b_{12} b_{13} b_{14}

b_{21} b_{22} b_{23} b_{24}

b_{25} b_{31} b_{32} b_{33}

b_{41} b_{42} b_{43} b_{44}

b_{45} b_{46} b_{47}

Figure 5. 3: Activity categories of tower crane.

Figures 5.2 and 5.3 showcase the action categories for each crane type, using examples from our dataset. These visuals highlight the practical application of our dataset in recognizing and classifying crane operations. To isolate specific actions, we used “Adobe Premiere Pro” to extract relevant video segments, which we then organized by crane activity. A Python script utilizing OpenCV was employed to automate frame extraction at a consistent rate, typically 1 frame per second (fps). For shorter videos, a 2-fps rate was used to ensure adequate representation. This approach provided equal samples for each category, facilitating unbiased training of our cascaded model.

5.2.2.2 Dataset preparation

After selecting the relevant video segments, we prepared the data to ensure consistency for model training and validation. We standardized the resolution and frame rate to create a uniform baseline, which helped eliminate variations from different video sources. Using “Roboflow,” we annotated the frames with bounding boxes and classified actions as safe, unsafe, near-miss incidents, and incidents according to a predefined taxonomy specific to crane operations. Our final dataset comprised 34,694 images depicting various scenarios. Through data augmentation techniques, we expanded this to 208,164 images, as detailed in Table 5.2. The dataset was then divided into training (80%), validation (10%), and testing (10%) sets to optimize the model's performance and development.

Table 5. 2: Final and augmented dataset preparation.

Equipment	Classes	Safety status	Subclasses	Original frames	Augmented frames	Final dataset
Mobile crane	A	Safe operations	a_{11}	1,649	8,245	9,894
			a_{12}	1,165	5,825	6,990
			a_{13}	2,740	13,700	16,440
			a_{14}	754	3,770	4,524
		Unsafe operations	a_{21}	813	4,065	4,878
			a_{22}	752	3,760	4,512
			a_{23}	541	2,705	3,246

			a_{24}	536	2,680	3,216
			a_{25}	957	4,785	5,742
		Near-miss incidents	a_{31}	542	2,710	3,252
			a_{32}	766	3,830	4,596
			a_{33}	454	2,270	2,724
		Incidents	a_{41}	609	3,045	3,654
			a_{42}	2,233	11,165	13,398
			a_{43}	794	3,970	4,764
			a_{44}	506	2,530	3,036
			a_{45}	1,917	9,585	11,502
			a_{46}	642	3,210	3,852
			a_{47}	748	3,740	4,488
		Safe operations	b_{11}	1,203	6,015	7,218
			b_{12}	817	4,085	4,902
			b_{13}	1,554	7,770	9,324
			b_{14}	653	3,265	3,918
		Unsafe operations	b_{21}	716	3,580	4,296
			b_{22}	893	4,465	5,358
			b_{23}	405	2,025	2,430
			b_{24}	728	3,640	4,368
			b_{25}	679	3,395	4,074
Tower crane	B	Near-miss incidents	b_{31}	601	3,005	3,606
			b_{32}	471	2,355	2,826
			b_{33}	648	3,240	3,888
		Incidents	b_{41}	572	2,860	3,432
			b_{42}	1,409	7,045	8,454
			b_{43}	892	4,460	5,352
			b_{44}	591	2,955	3,546
			b_{45}	1,659	8,295	9,954
			b_{46}	522	2,610	3,132
			b_{47}	563	2,815	3,378
Total				34,694	173,470	208,164

5.2.2.3 Data processing

Before deploying our cascaded model, which included SRGAN, RT-DETR-L, DINOv2, and ViT, we performed critical data processing to ensure that the input data were of high quality

and correctly formatted. We prepared two optimized datasets for the detection and classification tasks.

For the detection dataset, we began with 34,694 images. The SRGAN model requires both low-resolution and high-resolution versions of datasets. Using the PIL library, we generated a low-resolution dataset by applying effects such as night, blur, noise, and low-resolution to the original images and resizing them to 64x64 pixels. The high-resolution dataset consists of original frames resized to 224x224 pixels. The SRGAN then generated super-resolution images of 512x512 pixels, which were normalized using predefined mean and standard deviation values. We annotated these images using "Roboflow" to label and identify the mobile and tower cranes.

The classification dataset (208,164) included super-resolution images combined with augmented data. We employed a variety of augmentation techniques to increase the dataset diversity, such as flipping, rotation, brightness, contrast adjustments, modifications in saturation and hue, affine transformations, grayscale conversion, Gaussian blurring, sharpness adjustments, auto-contrast equalization, and perspective transformation. These augmentations, implemented using PIL and PyTorch's torchvision transforms, aimed to broaden the model's learning conditions to improve its generalizability. Annotations, including bounding boxes and action categories, were converted into a machine-readable vector format for the RT-DETR-L model. The processed images were organized into batches suitable for the input requirements of the model.

5.2.3 Image pre-processing

5.2.3.1 SRGAN for image resolution enhancement

The proposed cascaded model utilizes a Super-Resolution Generative Adversarial Network (SRGAN) designed to upscale the quality of low-resolution images captured by construction

site surveillance systems (Ledig et al., 2016). The SRGAN framework uses advanced deep learning techniques to transform low-resolution images into super-resolution images while maintaining key textural details and overall image fidelity. Although various methods are available for image pre-processing, SRGAN is preferred due to the typically blurred and low-resolution nature of CCTV footage.

As illustrated in Figure 5.4, the SRGAN consists of two primary networks: the generator and the discriminator network. The generator utilizes a ResNet-like architecture and is tasked with enhancing the resolution of low-resolution images. It starts with a 9x9 convolutional layer and is followed by multiple 3x3 convolutional layers, incorporating batch normalization and Parametric Rectified Linear Unit (PReLU) activation functions. Residual blocks were used to fine-tune the image details and progressively upscale the resolution. PixelShuffle operations facilitate a 4x upscaling factor through sub-pixel convolution, ensuring that the content integrity is preserved during enhancement. The generator produced an image with a resolution of 256x256 pixels.

The discriminator, a sophisticated convolutional neural network (CNN), differentiates between the super-resolution images generated by the model and ground-truth high-resolution images. It begins with a 3x3 convolutional layer and progressively deepens through additional layers, each equipped with batch normalization. The architecture culminates in a fully connected layer and final neuron that evaluates the authenticity of the image. A pre-trained Visual Geometry Group-19 (VGG-19) model was integrated to boost the discriminator's feature extraction capabilities and align the training process with the perceptual quality metrics.

The training of the SRGAN involved 200 epochs in an adversarial setup. The generator produces super-resolution images during this process, whereas the discriminator distinguishes them from ground-truth high-resolution images. The content loss function ensures that the

generated images retain the key features of ground-truth high-resolution images. The training process iterates over each data batch, with losses for the generator and discriminator calculated and averaged per epoch. Model performance was validated without updating parameters, and the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) were used to gauge image quality in relation to the original images. Model checkpoints are saved based on the validation loss, and the best-performing model is retained. The SRGAN loss function integrates Binary Cross-Entropy (BCE) for adversarial performance, Mean Squared Error (MSE) for content fidelity, and VGG-based perceptual loss to preserve texture and quality. Algorithm 5-1 provides a detailed outline of the SRGAN-based image resolution enhancement approach.

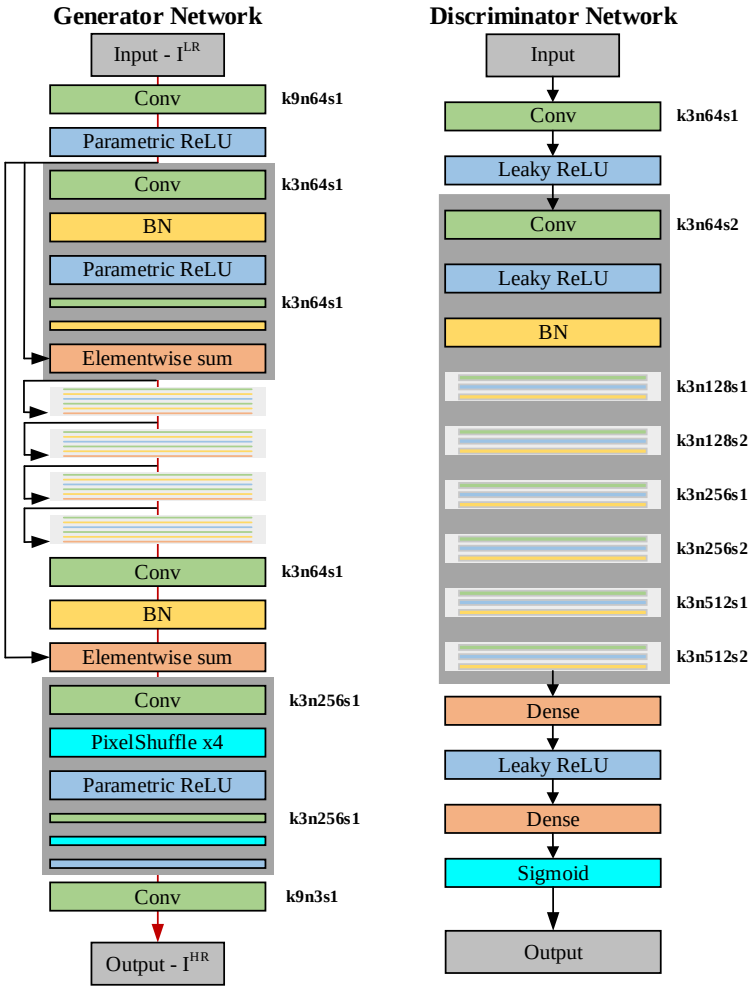


Figure 5. 4: SRGAN architecture.

Algorithm 5-1

SRGAN Algorithm

Input:

- $F_{LR} = \{f_{1,LR}, f_{2,LR}, \ldots, f_{n,LR}\}$ // A set of low-resolution frames
- $F_{HR} = \{f_{1,HR}, f_{2,HR}, \ldots, f_{n,HR}\}$ // A set of high-resolution frames

Output:

- $F_{SR} = \{f_{1,SR}, f_{2,SR}, \ldots, f_{n,SR}\}$ // A set of super-resolution frames

Method:

1. Initialize the generator *G* with weights ω_G
2. Initialize the discriminator *D* with weights ω_D

Training of *G*:

FOR each image pair $(l_r, h_r) \in (F_{LR}, F_{HR})$ **DO:**

- $sr \leftarrow G(l_r \mid \omega_G)$ // Generate SR from LR
- $L_{content} \leftarrow$ content_loss(sr, hr) // Calculate content loss
- $\omega_G \leftarrow \omega_G - \eta \nabla_{\omega_G} L_{content}$ // Updates weights using gradient descent

Adversarial Training:

REPEAT until convergence **DO:**

- **FOR** each $l_r \in F_{LR}$ **DO:**
  - $sr \leftarrow G((l_r \mid \omega_G)$ // Upscale LR to SR
  - $L_{content} \leftarrow$ content_loss(sr, HR) // Calculate content loss
  - $L_G \leftarrow$ adversarial_loss$(D(sr|\omega_D))$ // Adversarial loss for G
  - $\omega_G \leftarrow \omega_G - \eta \nabla_{\omega_G}(L_{content} + \lambda L_G)$ // Update G's weights
- **FOR** each image pair $(s_r, h_r) \in (F_{SR}, F_{HR})$ **DO:**
  - $L_{D_{real}} \leftarrow$ classification_loss$(D(hr|\omega_D)$, "real") // Loss for real hr
  - $L_{D_{fake}} \leftarrow$ classification_loss$(D(sr|\omega_D)$, "fake") // Loss for fake sr
  - $\omega_D \leftarrow \omega_D + \eta \nabla_{\omega_D}(L_{D_{real}} + L_{D_{fake}})$ // Update D's weights

Output:

RETURN $F_{SR} = \{f_{1,SR}, f_{2,SR}, \ldots, f_{n,SR}\}$ using the trained generator $G(sr|\omega_G)$.

5.2.4 Risk identification

In the risk identification phase, recognizing and listing potential crane operation hazards requires precise crane detection and hierarchical classification of operational activities. To achieve reliable detection and classification, dedicated models must be employed for each task within the cascaded pipeline. Using a single model can compromise efficiency, accuracy,

reliability, and adaptability. Therefore, we adopted cascade learning for automated safety risk assessment in crane operations.

Fig 5.5 shows the pipeline models for risk identification and detection in crane operations, addressing the challenges in low-quality CCTV footage. The pipeline combines three transformer-based models for risk identification, RT-DETR-L, DINOv2, and ViT, each serving distinct functions in the assessment framework. The cascaded pipeline begins with SRGAN, which enhances images through super-resolution reconstruction. Section 5.2.3 explains how this pre-processing step converts low-resolution surveillance imagery into high-fidelity images by reconstructing details and improving visual clarity. The enhanced images feed into RT-DETR-L, which detects mobile and tower cranes across frames. The RT-DETR-L temporal analysis enables object tracking through motion pattern analysis and temporal correlation. DINOv2 then performs primary activity classification, categorizing operations as safe, unsafe, near-miss incidents, or incidents. This layer establishes the basic safety status assessment. Finally, ViT performs detailed activity classification, characterizing specific operational actions. This classification enhances the precision of the safety assessment framework.

The integration of these transformer-based models creates a hierarchical pipeline that optimizes detection accuracy and computational efficiency for low-quality surveillance data. This architecture aims to achieve superior real-time performance by leveraging the specialized capabilities of each model. The framework shows potential for broad surveillance applications while addressing critical challenges in surveillance analysis, including low-quality footage processing, object tracking, and multitiered activity classification.

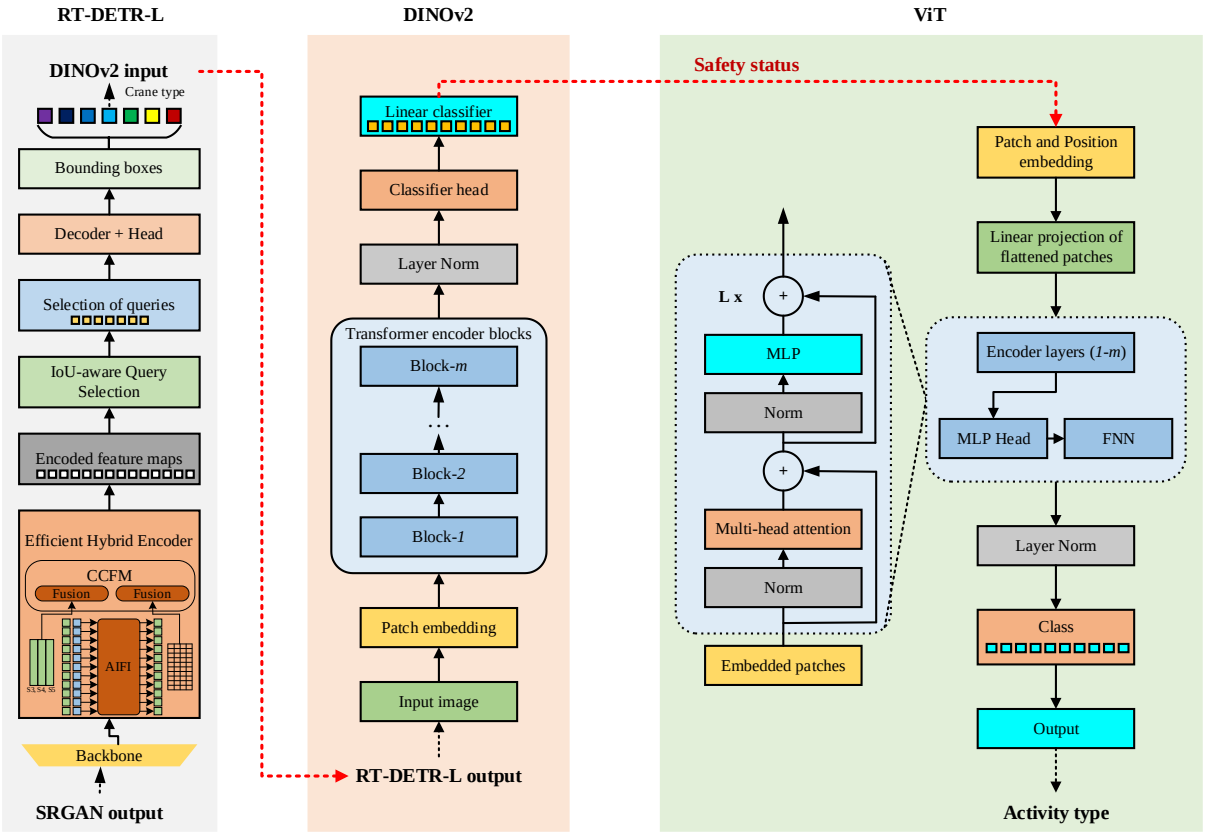


Figure 5. 5: RT-DETR-L, DINOv2, and ViT architectures.

5.2.4.1 RT-DETR-L for crane detection

The Real-Time Detection Transformer-Large (RT-DETR-L) is a sophisticated object detection model tailored for dynamic environments like construction sites (Carion et al., 2020). It enhances the DEtection TRansformer (DETR) framework, which treats object detection as a direct set prediction task, thereby eliminating the need for Non-Maximum Suppression (NMS) (Zhao et al., 2023). RT-DETR-L employs bipartite matching loss to align the predicted bounding boxes with ground-truth objects, simplifying detection and improving efficiency.

As illustrated in Figure 5-5, the RT-DETR-L architecture features an Efficient Hybrid Encoder, that integrates Attention-based Intra-scale Feature Interaction (AIFI) and a CNN-based Cross-scale Feature-fusion Module (CCFM) to handle multi-scale features effectively. To enhance the accuracy of crane detection and classification (specifically for mobile and tower cranes), RT-DETR-L was coupled with the SRGAN. This integration boosts the image resolution, which

in turn improves detection accuracy. The model begins with HGStem, which increases the channel depth while maintaining data integrity, followed by hierarchical grouped blocks (HGBlock) that refine feature maps and optimize spatial resolution. The responsiveness of the model to object motion and temporal variations in dynamic scenes is further enhanced by Intersection over Union (IoU)-aware Query Selection and temporal embeddings from previous frames.

The RT-DETR-L model underwent training for 10 epochs using a dataset comprising 34,694 super-resolution annotated images. The training employed the AdamW optimizer and utilized loss functions such as generalized intersection over union (GIoU), Cross-Entropy, and L1 loss to enhance the detection, classification, and localization precision. The integration with SRGAN enables RT-DETR-L to process upscaled images, thereby markedly improving its detection performance. Validation conducted on 3,469 images revealed significant improvements in the accuracy metrics, including mAP50 and mAP50-95. These findings underscore the effectiveness of RT-DETR-L for real-time crane detection and classification within dynamic construction environments, as detailed in Algorithm 5-2.

Algorithm 5-2

RT-DETR-L Algorithm

Input:

- $F = \{f_1, f_2, \ldots, f_n\}$ // A set of annotated frames from SRGAN
- $E = \{A, B\}$ // A set of construction equipment; A = Mobile crane and B = Tower crane

Output:

- $AF = \{af_1, af_2, \ldots, af_n\}$ // A set of annotated frames with detected objects A and B

Method:

1. Initialize RT-DETR-L model (M) with weight ω:

Temporal Feature Extraction:

FOR each frame $f_i \in F$ **DO:**

- $T_f \leftarrow$ extract_temporal_features $(f|\omega)$ // Extract temporal features from frame

Object Detection and Annotation:

FOR each temporal feature T_f in sequence **DO:**

- AB ← detect_object (T_f | ω, E) // Detect objects A and B for classes
- AF_f ←annotate_frame (f, AB) // Annotate f_i with detected object AB

Real-Time Optimization:

WHILE processing video stream **DO:**

- **FOR** each new incoming frame f_{new} **DO:**
  - Update T with features from f_{new} considering previous frames
  - AB_{new} ← detect_objects (T / ω, E)
  - $AF_{f_{new}}$ ← annotate_frame (f_{new}, AB_{new})
  - ω ← adapt weights (ω, AB_{new})

Output:

RETURN the sequence of annotated frames $AF = \{af_1, af_2, \ldots, af_n\}$ using the trained RT-DETR-L.

5.2.4.2 DINOv2 for safety classification

DINOv2 is a state-of-the-art self-supervised vision transformer model introduced by Oquab et al. (2023), building upon the foundation of its predecessor, DINO. Unlike traditional computer vision models that heavily depend on labeled data, DINOv2 learns robust visual features without requiring manual annotations. This model addresses the limitations of conventional approaches by capturing contextual information that is often missed in manually labeled descriptions. The core concept behind DINOv2 is self-distillation, which is a process in which the model learns to predict its own output when presented with slightly altered versions of the same image. This method has enabled DINOv2 to develop versatile visual features that perform well across various computer vision tasks such as classification, segmentation, and image retrieval, even without task-specific fine-tuning. Notably, DINOv2 has demonstrated strong capabilities in challenging areas such as direct depth estimation from images.

In our study, we adapted the DINOv2 framework to classify crane operations as safe, unsafe, near-miss incidents or incidents, demonstrating the model's adaptability and potential for diverse applications. As shown in Figure 5-5, the architecture of DINOv2 is based on a Vision Transformer (ViT) backbone and employs a linear classifier for self-supervised feature

extraction. The model processes the input images by dividing them into a grid of non-overlapping patches, each projected into a high-dimensional vector space. This allows for an in-depth analysis of local features within patches, whereas the self-attention mechanism integrates these details into a comprehensive understanding of the context of the image. The model consists of 12 transformer blocks, each equipped with layer normalization and Memory-Efficient Attention mechanisms to enhance its ability to discern the relationships between different image elements.

During training, the input images were resized to a uniform 512x512 dimension and normalized. The model was trained and validated using a classification dataset comprising 208,164 samples to ensure generalization ability. We employed an 80-10-10 split for the training, validation, and test sets, with the model undergoing 10 training epochs. The cross-entropy loss function and the Adam optimizer were used for the optimization. The validation dataset was utilized to monitor the performance and make necessary adjustments, offering insights into the effectiveness of the model on unseen data. DINOv2 was integrated with other models, such as SRGAN and RT-DETR-L, in the proposed approach to create a cascaded system. SRGAN first enhances low-resolution surveillance images, which RT-DETR-L then processes to identify and localize cranes, including mobile and tower cranes. DINOv2 subsequently classifies these localized objects to monitor operational safety. This end-to-end pipeline ensures a streamlined workflow in which each model's output directly informs the next, resulting in accurate real-time safety assessments. Algorithm 5-3 of DINOv2 provides a real-time classification of crane activities into safe, unsafe, near-miss incidents, and incidents categories.

Algorithm 5-3

DINOv2 Algorithm
Input:

- $F = \{f_1, f_2, \ldots, f_n\}$ // A set of frames
- $S = \{Safe, Unsafe, near\ miss\ incidents, incidents\}$ // Safety status

Output:

- $AF = \{af_1, af_2, \ldots, af_n\}$ // A set of annotated frames with safety status

Method:

1. Initialize DINOv2 model (M) with weights ω
2. Define resizing dimensions and patch size

Safety Status Classification:

3. **FOR** each frame $f \in F$ **DO:**
    - $f_{normalized} \leftarrow$ resize and normalized (f)
    - $\{p_i\} \leftarrow$ split_into_patches ($f_{normalized}$)
    - $\{v_i\} \leftarrow$ project_patches ($\{p_i\}, M$)
    - $T_f \leftarrow$ apply_transformers ($\{v_i\}, M$)
    - $O_f \leftarrow$ classify_objects (T_f, S) // Classify objects in terms of safe, unsafe, near-miss incidents, and incidents
    - $AF_f \leftarrow$ annotate_frame (f, O_f)
    - Update ω
        - $\omega \leftarrow$ adapt_weights (ω, O_f)

Output:

- **RETURN** $AF = \{af_1, af_2, \ldots, af_n\}$ using the trained DINOv2 model.

5.2.4.3 ViT for activity classification

The Vision Transformer (ViT) introduced by Dosovitskiy et al. (2020) represents a significant advancement in image recognition and classification. By dividing the images into smaller patches and capturing both local and global relationships, ViT addresses the limitations of earlier attention mechanisms. This method enables the model to exponentially expand its receptive fields without adding an additional computational burden. ViT demonstrated exceptional performance in image classification tasks, particularly when trained on large datasets, surpassing state-of-the-art CNN models in terms of both accuracy and computational efficiency.

As illustrated in Figure 5-5, the ViT model begins its process with patch and position embedding. The image is divided into fixed-size square patches, each of which is linearly

transformed into a vector. Position encodings are then added to these vectors to retain the spatial information. The model is composed of multiple encoder layers, each containing two sub-layers: multi-head self-attention and feedforward neural networks (FNN). The multi-head self-attention mechanism computes the weighted relevance between patches, allowing the model to capture the relationships between different parts of the image. The resulting output is processed by an FNN, often with an activation function such as a Rectified Linear Unit (ReLU), which introduces nonlinearity and enables the model to learn more complex patterns. Layer normalization was applied to stabilize and accelerate the training process by normalizing the inputs to each sublayer. Additionally, residual connections were used to add the original input embeddings to the output of each sublayer, ensuring the preservation of information across layers.

Training ViT models involves several steps, including dataset processing and normalization, batch division, and the use of a training loop. The cross-entropy loss function and the Adam optimizer were employed to update the model weights over multiple epochs. Hyperparameters such as the learning rate and batch size were carefully selected to balance the speed and stability of the training process. Performance metrics were used to evaluate the ability of the model to accurately classify and segment image regions. The validation of unseen data is crucial for assessing the model's generalization capabilities, identifying potential overfitting, and informing adjustments to hyperparameters. The objective was to balance low validation loss and high accuracy, indicating a well-generalized model with strong predictive capabilities. The ViT algorithm, as described in Algorithm 5-4, facilitates the real-time classification of activities.

Algorithm 5-4

ViT Algorithm
Input:

- $F = \{f_1, f_2, \ldots, f_n\}$ // A set of frames
- $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$ // Set of class labels

Output:

- $C = \{c_1, c_2, \ldots, c_n\}$ // A set of predicted class

Method:

1. Initialize ViT model (M) with weights ω
2. Define patch size, embedding size, and image size.

Activity classification:

3. **FOR** each image $f_i \in F$ **DO:**

- $f_{normalized} \leftarrow$ resize and normalized(f)
- $\{p_j\} \leftarrow$ split_into_patches($f_{normalized}$, patch_size)
- $\{v_j\} \leftarrow$ project_patches($\{p_j\}$, M)
- $T_i \leftarrow$ apply_transformers($\{v_j\}$, M)
- $O_i \leftarrow$ classification_head(T_i, C) // Obtain class probabilities
- $p_i \leftarrow O_i$ // Store the class probabilities

Output:

- **RETURN:** $C = \{c_1, c_2, \ldots, c_n\}$ using the trained ViT model

---

5.2.5 Risk assessment and analysis

After identifying potential crane operation hazards, the next step is to evaluate the impact and likelihood of each risk. Risk assessment consists of two components: (1) quantifying the initial risk values through confidence levels and probabilities and (2) applying a quantitative method to determine risk levels. The severity assessment phase categorizes potential hazardous consequences into four groups: safe, unsafe, near-miss incidents, and incidents. Hazard probability was estimated using confidence levels. Each activity received a risk value based on its confidence level, probability, and impact severity. To prioritize urgent hazards, a risk matrix combines the severity and likelihood ratings using two concepts: probability-derived threshold values and severity levels established by Hou et al. (2023). We adopted their risk-scoring system, which was developed through expert questionnaires, to optimize the assessment process. Table 5-3 presents the detailed risk levels.

Table 5. 3: Risk levels.

S. No	Safety status	Risk level	Description for consequences	Score
1	Safe operations	I	Negligible consequences	0-20
2	Unsafe operations	II	Mild consequences	21-40
3	Near-miss incidents	III	Moderate consequences	41-60
4	Incidents	IV	Severe consequences	61-100

5.2.6 Risk evaluation

In the risk evaluation, we prioritize risks based on their severity and likelihood to determine which activities require immediate action. For this purpose, we establish threshold values using two criteria: (1) a probability-based threshold and (2) a risk-scoring system. These predefined thresholds enable operators and decision-makers to monitor crane operations effectively. If the value of an activity falls below the predefined threshold, it indicates that the crane is operating in a safe environment. However, if the comprehensive risk value of an activity exceeds a predefined threshold, immediate mitigation and control measures are required.

5.2.7 Control and mitigation

This step implements strategies to eliminate or reduce identified risks to acceptable levels by sending real-time warning signals to operators. Operators can view images of each activity along with their comprehensive risk values and warning signs. This enables operators to stop work or implement alternative solutions to reduce potential risks.

5.2.8 Monitoring and documentation

In this step, we continuously monitor crane operations and review risk management practices to ensure their effectiveness. We track crane operations through video surveillance to detect new hazards. The safety data collected via cascaded pipeline helps determine whether control measures are working effectively or whether new risks have emerged. We conduct periodic

audits and update risk assessments based on findings. Additionally, we maintain detailed records of all risk assessment processes, findings, and implemented actions.

5.3 Cascade pipeline for safety risk assessment

The cascaded model integrates four distinct models: SRGAN, RT-DETR-L, DINOv2, and ViT for automated safety risk assessment of crane operations in noisy and low-quality CCTV footage from construction sites. Each model serves a specific role: 1) SRGAN enhances image resolution; 2) RT-DETR-L detects mobile and tower cranes; 3) DINOv2 performs the initial (Level-I) classification to assess crane operation safety; and 4) ViT conducts detailed (Level-II) classification to identify specific crane activities.

This model significantly enhances computational efficiency, accuracy, reliability, and adaptability in detecting and classifying unsafe crane operations within challenging surveillance footage. As illustrated in Figure 5-6, the process begins with SRGAN, which preprocesses low-resolution CCTV images by upscaling them to 512×512 pixels and applying Gaussian blur to reduce noise while preserving essential details. Subsequently, RT-DETR-L identifies cranes within these enhanced images using specialized training. DINOv2 then assesses the safety of each detected crane action by analyzing the context, usage, and surrounding environment, adding a critical safety evaluation layer that classifies crane operations into safe operations, unsafe operations, near-miss incidents, and incidents. ViT subsequently provides a detailed analysis of the classified crane actions, distinguishing specific activities based on an extensive dataset of crane operations. A comprehensive risk value is then determined based on confidence level and converted to probability. For severity and likelihood of risk, impact levels are incorporated with each activity. A predefined threshold value is established based on probability and impact scores. When the comprehensive risk value exceeds the predefined threshold, a real-time warning signal is sent to operators for immediate

risk mitigation actions. The process is continuously tracked across various additional videos, and records of data, processes, findings, and actions are maintained in the documentation step. The procedure for the cascaded transformer-based model is detailed in Algorithm 5-5.

This model demonstrates considerable potential for automating safety management in the construction industry. The approach offers a comprehensive view of potential hazards on construction sites by systematically enhancing image quality, detecting cranes, assessing safety, and identifying specific crane activities. These capabilities can contribute to the development of more effective safety protocols and improve overall safety outcomes in construction environments.

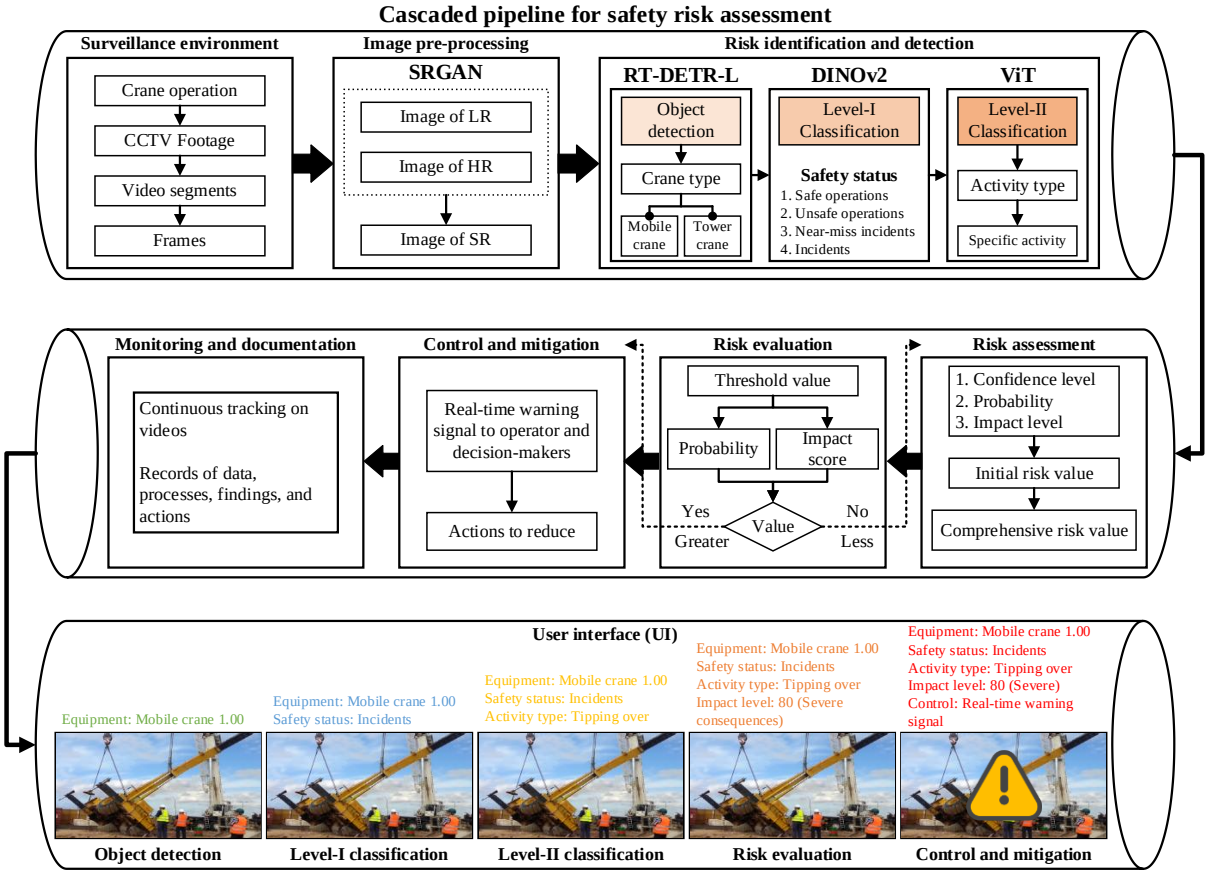


Figure 5. 6: Cascaded model architecture for automated safety risk assessment for crane operations.

Algorithm 5-5

Cascaded Model Algorithm

Input:

- $F = \{f_1, f_2, \ldots, f_n\}$ // A set of frames.
- $E = \{A, B\}$ // A set of construction equipment; A = Mobile crane and B = Tower cranes
- $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$ // A set of class labels
- $S = \{Safe\ operations, Unsafe\ operations, near\ miss\ incidents, incidents\}$ // Safety status

Output:

- $AF_A = \{af_{A,1}, af_{A,2}, \ldots, af_{A,n}\}$ and $AF_B = \{af_{B,1}, af_{B,2}, \ldots, af_{B,n}\}$ // A set of annotated frames with classified objects and activities in A and B.

Method:

1. **FOR** each f in F **DO**:
   - $\forall f_i \in F, f_{i,SR} \leftarrow SRGAN(f_i)$ // Enhance the resolution of the frame
2. **Initialize** the RT-DETR-L model M with weight ω
   - M ← RT-DETR-L(ω) // Initialize the model M

Object detection:

3. **FOR** each $f_{i,SR}$ in sequence **DO**:
   - $\forall f_{i,SR} \in \{f_{1,SR}, f_{2,SR}, \ldots, f_{3,SR}\}, E_{detected,i} \leftarrow detect\_objects(f_{i,SR} \mid \omega, E)$

Level-I Classification (Safety status):

4. **FOR** each detected object (AB) in $E_{detected}$ **DO**:
   - $\forall AB_j \in E_{detected,i},\ safety\_class_{j,i} \leftarrow DINOv2(AB_j \mid \omega)$

Level-II Classification (Activity type):

5. **IF** $(safety\_class_{j,i} = Unsafe)$ **THEN**
   - **FOR** each unsafe object AB_{Unsafe} in $E_{detected}$ **DO**:
     - $\forall AB_{unsafe,l} \in E_{detected,i},\ activity\_type_{l,i} \leftarrow ViT(ab_{unsafe,l} \mid \omega, A\&B)$

Risk assessment:

6. **FOR** each activity in $\mathbf{E}_{(detected,i)}$ **DO**:
   - ∀activity_k ∈ {activity_1, activity_2, ..., activity_m}:
     - risk_value_k ← P_k(confidence_level_k)

Risk evaluation:

7. **FOR** each risk_value_k **DO**:
   - ∀risk_value_k ∈ {risk_value_1, risk_value_2, ..., risk_value_m}:
     - **IF** (risk_value_k ≥ threshold_value) **THEN**
       - impact_level_k ← high

       ELSE:
       - impact_level_k ← low

Control and mitigation:

8. **FOR** each impact_level_k **DO**:
   - ∀impact_level_k ∈ {impact_level_1, impact_level_2, ..., impact_level_m}:
     - **IF** (impact_level_k = high) **THEN**

        - control_k ← implement_control(activity_k)

Monitoring and documentation:

9. **FOR** each f in F continuously **DO**:
    - ∀f_i ∈ {f_1, f_2, ..., f_n}:
        - record_operation(f_i)
    - monitor_new_videos()

Annotation and UI integration:

10. **FOR** each $f_{i,SR}$ with detected objects and classifications **DO**:
    - $\forall f_{i,SR} \in \{f_{1,SR}, f_{2,SR}, \ldots, f_{n,SR}\}$, $AF_i \leftarrow annotate\_frame(f_{i,SR}, E_{detected,i}, safety\_status_i, activity\_type_i)$
    - Update_UI(AF_i) // update the User Interface (UI) with the annotated frame

RETURN:

- **RETURN**: $F = \{AF_1, AF_2, \ldots, AF_n\}$

5.4 Results

5.4.1 Experimental setup

As detailed in Table 5-4, this study utilized advanced computational resources to handle the complex data processing and model training requirements. We utilized a high-performance computing (HPC) cluster known as the "University Research Facility in Big Data Analytics (UBDA)," which integrates cutting-edge hardware and software to support the intensive computational tasks involved in our experiments.

Table 5. 4: Experimental setup.

Component	Specification	Quantity
CPU	AMD EPYC9124	2
Cooler	Supermicro Heavy duty fans	8
RAM	DDR5 Samsung MultiBitECC, 64GB M321R8GA0BB0-CQKZJ	16
SSD	Samsung SSD 870 EVO 500GB	2
VGA	NVIDIA RTX6000Ada 48GB	8
Motherboard (M/B)	Supermicro H13DSG-O-CPU	1
Power supply	Supermicro PWS-2K08A-1R 2000W 80+ Platinum	4

5.4.2 Performance evaluation of models in a pipeline

This section presents a comprehensive performance evaluation of a cascaded pipeline integrating four distinct models: SRGAN, RT-DETR-L, DINOv2, and ViT. The integrated model was developed to enhance construction site safety, advance SSSS technologies, and refine safety protocols through improved object detection and hierarchical classification in automated safety risk assessment. The following subsections detail the experimental results, encompassing computational complexity, accuracy, precision, reliability, and adaptability metrics for each model, as well as their synergistic performance within the cascaded pipeline.

5.4.2.1 Resolution enhancement via SRGAN

CCTV footage ensures safety at construction sites. However, their low resolution often hinders effective object detection and classification in the risk identification stage. To overcome this challenge, we developed a novel cascaded model integrating SRGAN, RT-DETR-L, DINOv2, and ViT to enhance safety monitoring. A key component of our pipeline is the SRGAN model, which enhances the low-resolution CCTV footage by converting it into high-resolution images. This improvement in visual data quality facilitates more accurate detection and classification in the subsequent stages. The SRGAN is particularly effective at upscaling images while maintaining detail and synthesizing high-frequency information typically lost at lower resolutions, which is crucial for accurately capturing complex environments such as those during crane operation.

As shown in Figure 5-7, we performed a comparative analysis using before-and-after images to highlight the effectiveness of the SRGAN model. The visual analysis demonstrated significant improvements in image quality, making previously inadequate footage suitable for accurate evaluation. We used the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) to quantify these enhancements. PSNR assesses the ratio of the maximum possible signal power to the power of the noise that distorts the signal. Simultaneously, SSIM

evaluates the similarity between the ground-truth high-resolution images and super-resolution images produced by our model. These metrics were calculated using Equations (5-1) – (5-6):

Our model achieved an average PSNR of 15.49 dB and an SSIM of 0.87 on the test dataset, indicating a notable improvement in image quality. Higher PSNR values generally suggest lower image degradation, and SSIM values closer to 1 indicate a better quality. Our PSNR results surpassed those previously reported in the literature, including those of Zhang et al. (2024) (23.24 dB) and Kim et al. (2024) (35.57 dB). This enhancement lays a strong foundation for subsequent object detection and classification tasks, leading to more effective construction-site monitoring. Improved resolution is critical for accurately detecting and classifying safe operations, unsafe operations, near-miss incidents, and incidents, thereby significantly boosting the overall performance of our safety monitoring system.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{i=0}^{n-1}[f_{HR}(i,j) - f_{SR}(i,j)]^2 \tag{5-1}$$

$$PSNR = 10\ log\left(\frac{MAX^2}{MSE}\right) \tag{5-2}$$

$$SSIM = \frac{(2\mu_{HR}\mu_{SR} + (0.01L)^2(2\sigma_{HR,SR} + (0.03L)^2}{(\mu_{HR}^2 + \mu_{SR}^2 + (0.01L)^2)(\sigma_{HR}^2 + \sigma_{SR}^2 + (0.03L)^2)} \tag{5-3}$$

$$precision = \frac{TP}{TP + FP} \tag{5-4}$$

$$recall = \frac{TP}{TP + FN} \tag{5-5}$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} p(r)dr \tag{5-6}$$

In these equations: $f_{HR}(i,j)$ and $f_{SR}(i,j)$ represent the ground-truth high-resolution and generated super-resolution frames, respectively. The variables m and n denote the dimensions of the image, and MAX refers to the maximum pixel value. The symbols μ_{HR} and μ_{SR} indicate the mean pixel values of ground-truth and generated images, while σ_{HR} and σ_{SR} denote the

variances of the ground-truth and generated images. The term $\sigma_{HR,SR}$ describes the covariance between the original and generated images, and L is the pixel value range. Additionally, TP, FP, and FN indicate true positives, false positives, and false negatives, respectively, whereas p and r stand for precision and recall.

Figure 5. 7: Example of resolution enhancement: (a) – (h) original and (a1) – (h1) generated images.

5.4.3 Risk identification

5.4.3.1 Crane detection via RT-DETR-L

The increasing complexity of construction sites and the extensive use of CCTV for remote surveillance necessitate sophisticated object detection solutions. To meet this demand, a cascaded model was designed that incorporates super-resolution techniques with cutting-edge transformer-based object detection and hierarchical classification models.

The process begins by enhancing the low-resolution CCTV footage using the SRGAN model. This initial step provides a more precise and detailed input for the subsequent transformer-based RT-DETR-L model specifically designed to identify and locate mobile and tower cranes in these challenging environments. RT-DETR-L excels in understanding the spatial relationships between elements in a scene and can process sequences of frames to track objects over time. These capabilities are valuable in the ever-changing conditions of construction sites, where rapid adaptation to evolving circumstances is crucial.

As shown in Figure 5-8, the model demonstrated an impressive performance in accurately detecting and localizing cranes in high-resolution images. This precision is essential in cluttered construction environments, where cranes may be partially obscured or blended into complex backgrounds. The output of the model typically includes bounding boxes around the detected cranes, along with labels and confidence scores for each detection. For instance, as depicted in Figure 5-8(a), the detected object was a mobile crane with a confidence score of 96.00%. The overall detection precision of our model was 92.10%, which surpasses the detection precision reported in the literature, such as 90.26% by Shin et al. (2024) and 90.26% by Lee et al. (2022). The preliminary results from the RT-DETR-L model show great promise for handling complex construction environments and challenging conditions. This precision is crucial for monitoring crane operations and personnel positioning, potentially enhancing overall site safety.

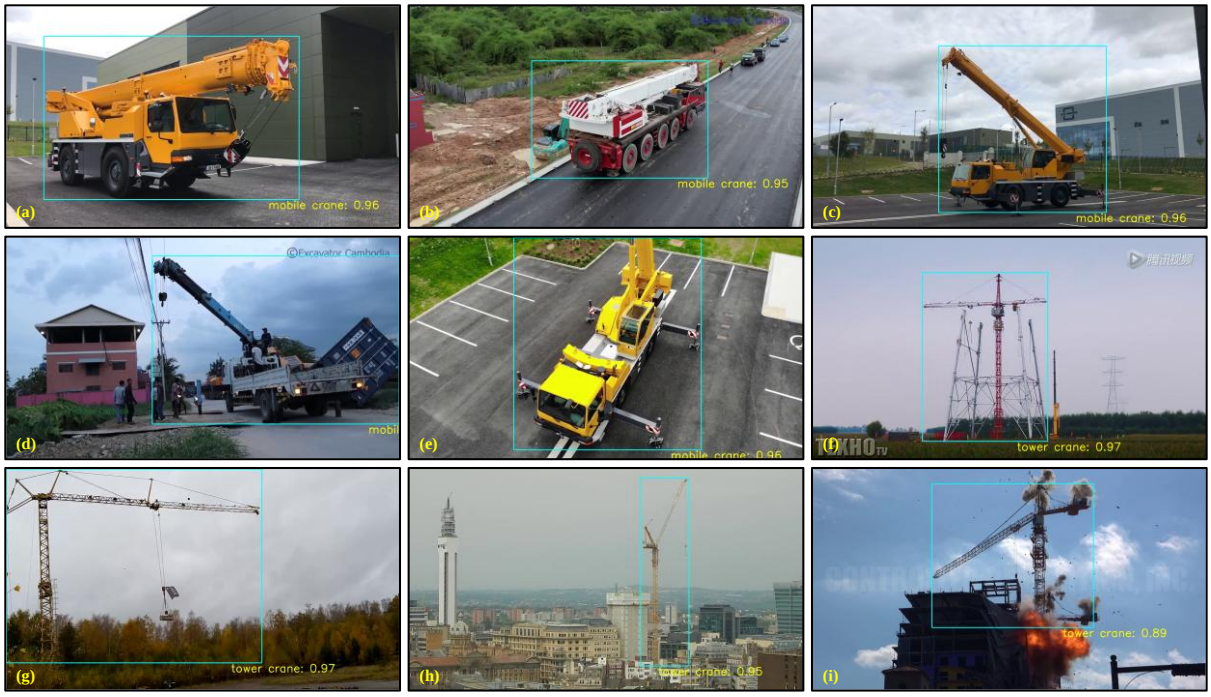


Figure 5. 8: Crane detection: (a) – (e) Mobile crane and (f) – (i) Tower crane.

5.4.3.2 Safety status via DINOv2 (Level-I classification)

The hierarchical classification of detected cranes is essential for understanding and categorizing crane operations. This classification consists of two levels: Level-I classifies the safety status of operations, whereas Level-II identifies the specific activity types. This approach provides deeper insights into the nature of the operations performed by mobile and tower cranes.

For Level-I classification, we employed a transformer-based DINOv2 model to differentiate between safe operations, unsafe operations, near-miss incidents, and incidents, providing critical insights for safety monitoring. The performance of the model in categorizing activities into safety classes is illustrated using a confusion matrix, as shown in Figure 5-9. This matrix offers a clear visual representation of the ability of the model to distinguish between safe operations, unsafe operations, near-miss incidents, and incidents, which is essential for accident prevention. The confusion matrix revealed a high true-positive rate of 98.72% for safe operations, 99.42% for unsafe operations, 99.69% for near-miss incidents, and 99.26% for

incidents. These results, derived from evaluations of the test dataset, highlight the effectiveness of the model in handling new, unseen conditions. The DINOv2 model's high accuracy in classifying safe operations, unsafe operations, near-miss incidents, and incidents can be attributed to the quality of the training dataset and the model's ability to capture complex patterns and relationships within the data.

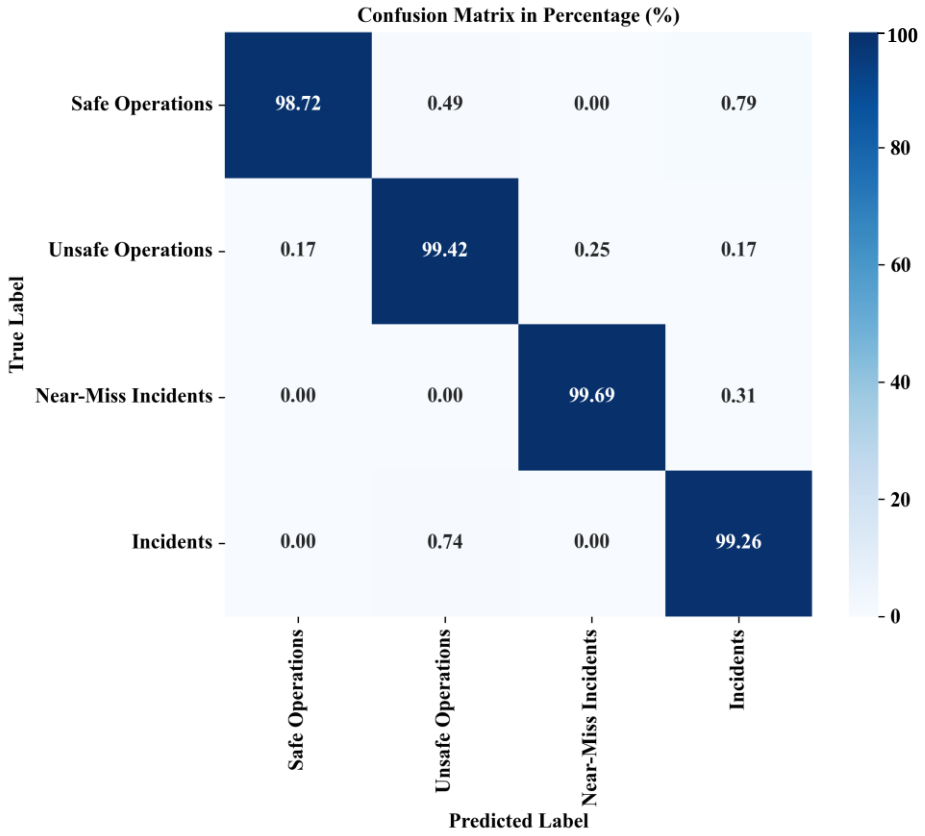
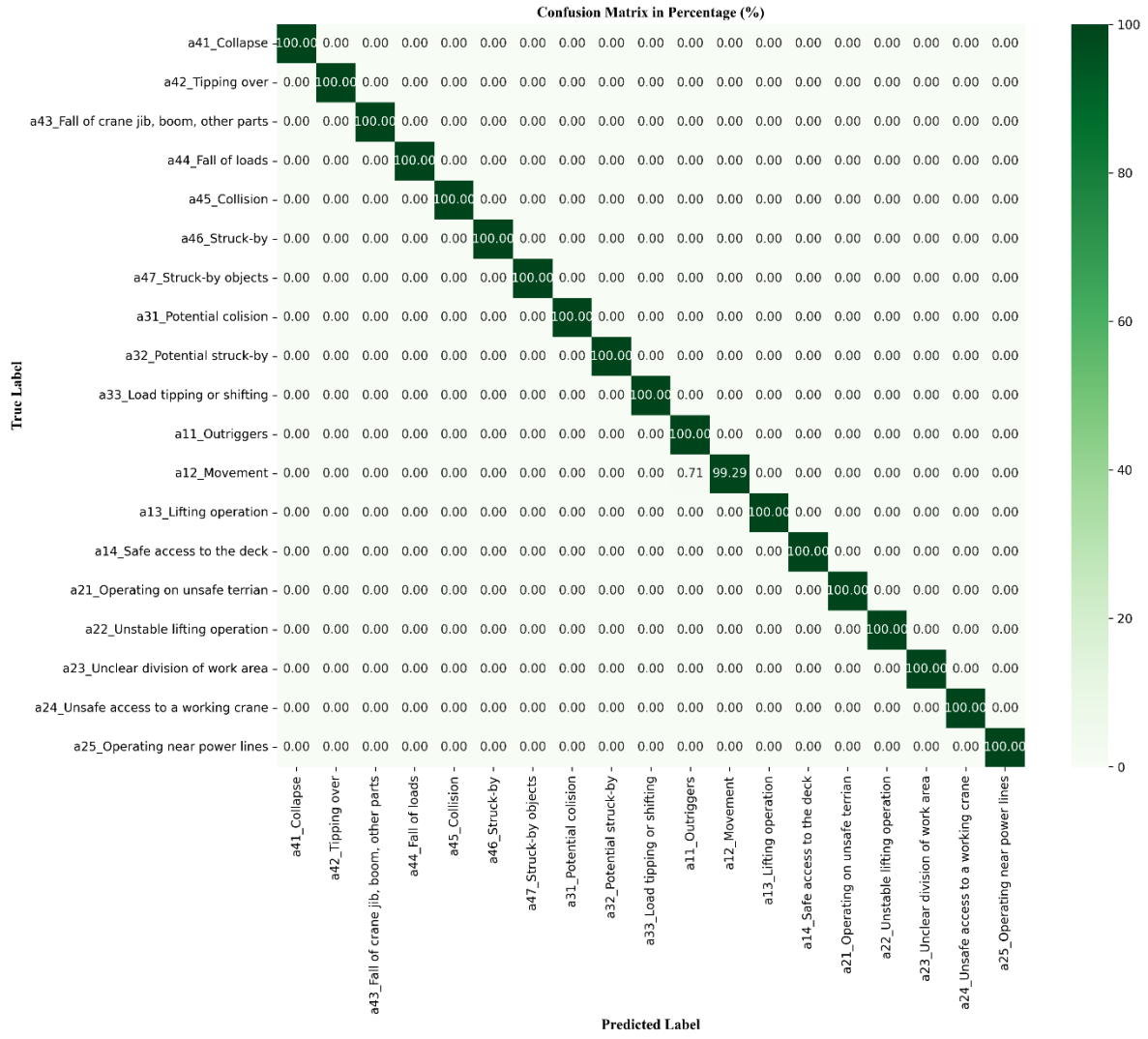


Figure 5. 9: Confusion matrix for Level-I classification.

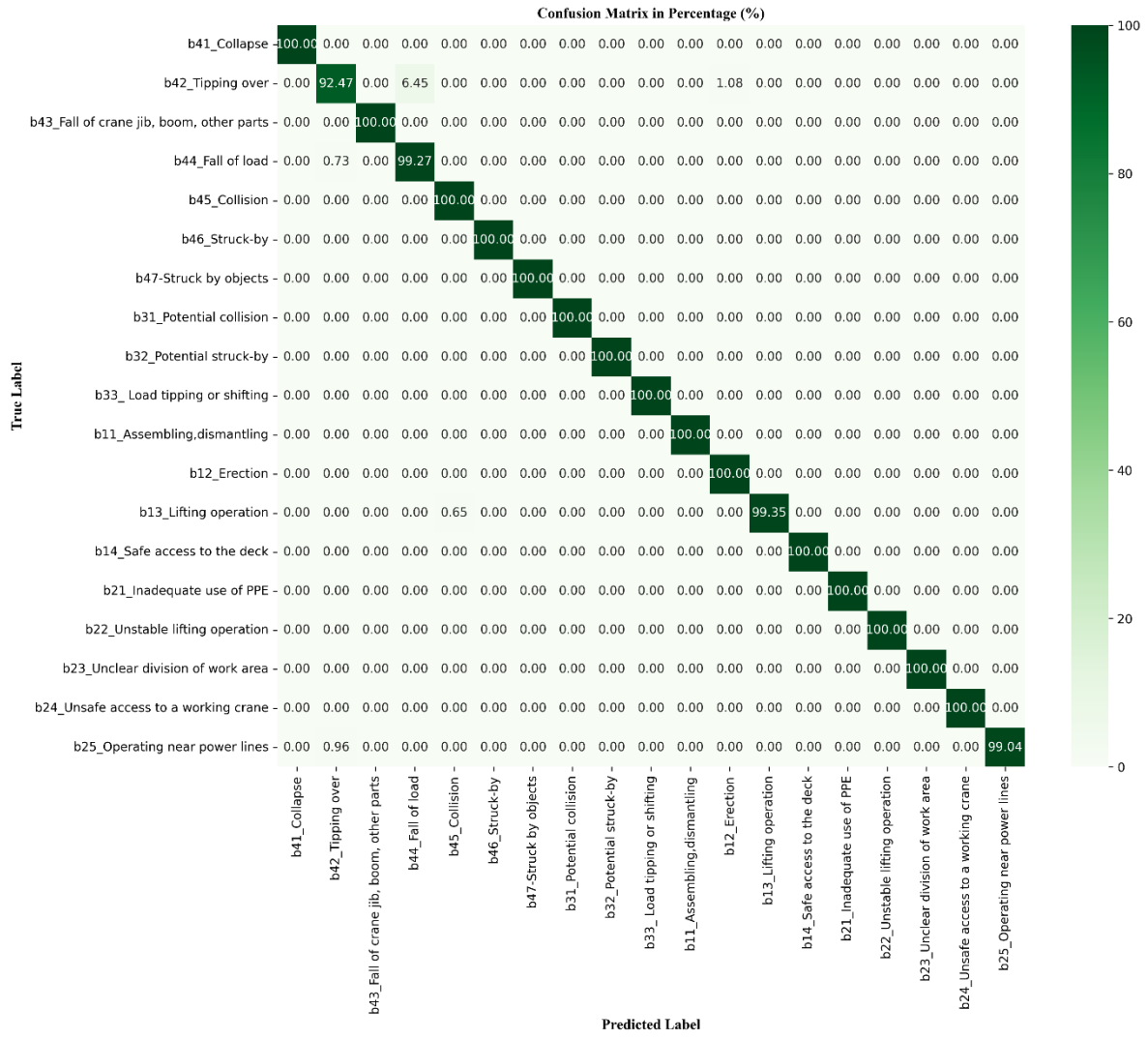
5.4.3.3 Activity recognition via ViT (Level-II classification)

For Level-II classification, a Vision Transformer (ViT) model was employed to categorize the specific activity types of mobile (100%) and tower (99.56%) cranes. The performance of the ViT model in classifying these activities is depicted in confusion matrices, as shown in Figure 5-10. These matrices illustrate the high accuracy of the model in sorting activities into predefined classes, demonstrating its effectiveness in handling new, unseen conditions. The diagonal values in the confusion matrices represent the proportion of correctly identified

activities, underscoring the model's sensitivity and specificity in distinguishing between the various activity types. The results indicated minimal false positives and false negatives, highlighting the practical effectiveness of the cascaded model in improving on-site safety protocols.



(a)



(b)

Figure 5. 10: Confusion matrices for Level-II classification: (a) Mobile crane and (b) Tower crane.

5.4.4 Risk assessment and evaluation

As discussed in Sections 3.5 and 3.6, the impact and likelihood of each activity were determined from confidence levels in probability. The severity and impact level were incorporated into each activity using Equation (5-7). Additionally, we utilized the risk level for each activity as outlined in Table 5-3. Following the assessment, predefined threshold values were established for both cases. Specifically, the probability threshold was set at 0.52, while the risk-scoring system threshold was set at 60. Activities exceeding these defined thresholds required control and mitigation measures.

$$r_i = \frac{\sum_{z=1}^{19}(a_{i,z} * I_{a_{i,z}})}{\sum I_{a_{i,z}}} \quad (5\text{-}7)$$

The comprehensive risk value (r_i) is determined for each activity (a_i), considering the corresponding impact level (I_i).

5.4.5. Control, monitoring, and documentation

In the control stage, operators receive visual signals indicating activity type, risk level, and safety status. For example, in Figure 5-11(a), the activity involves outriggers with a negligible risk level, indicating a safe environment. Conversely, in Figure 5-11(b), the activity displays a severe risk level with incident status, exceeding the threshold values of 0.52 or 60. Consequently, a real-time warning signal was transmitted to the operator, prompting immediate risk reduction or mitigation actions. The effectiveness of the cascade model for safety risk assessment was validated through extensive testing with diverse new video datasets, with comprehensive documentation of data, processes, findings, and implemented actions.

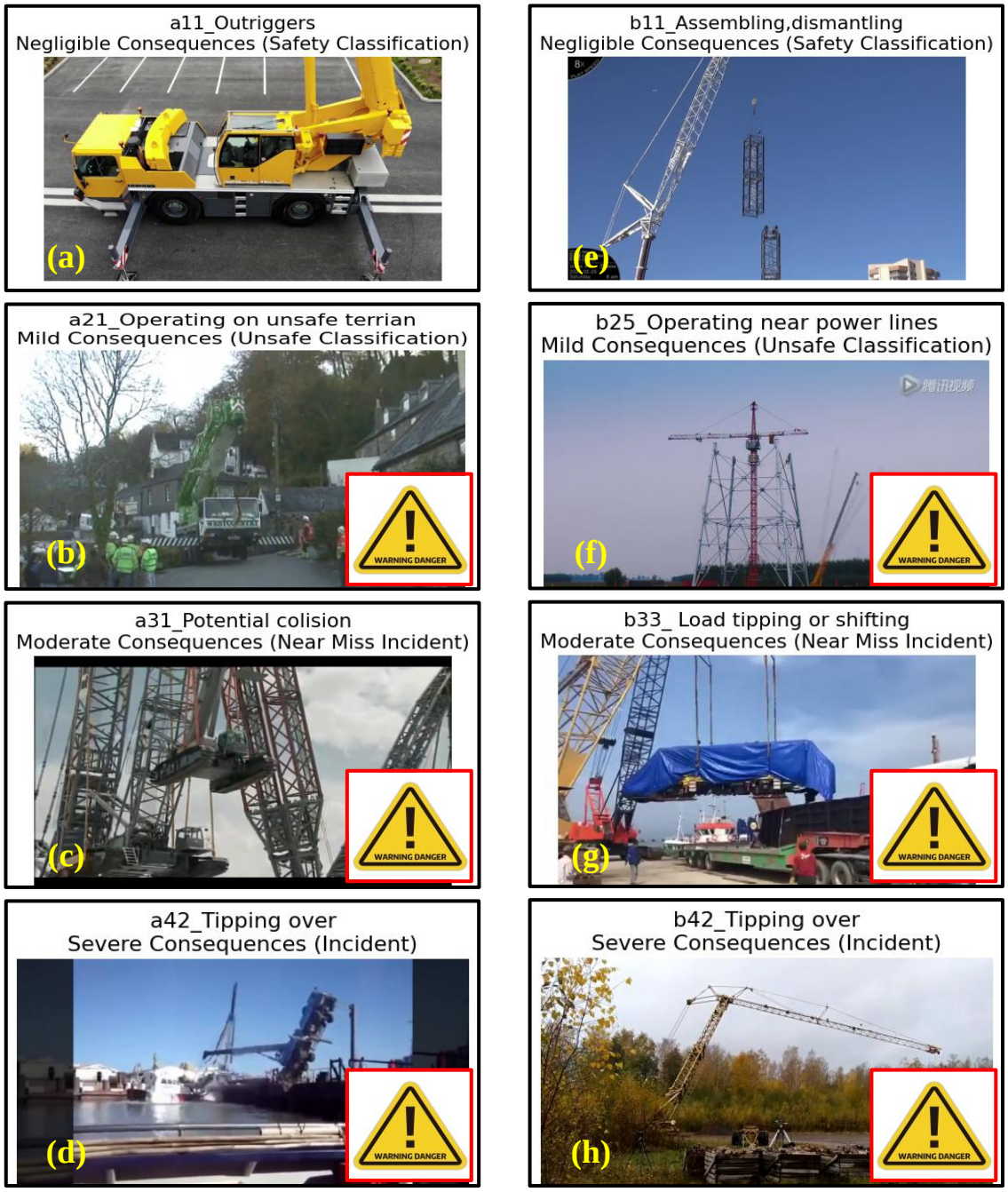


Figure 5. 11: Example of control and monitoring activities: (a) Mobile crane and (b) Tower crane.

5.5 Overall performance of a cascaded model

Our study assessed how well the cascaded model improves construction site safety by enhancing object detection and hierarchical classification in the risk identification stage. We examined various metrics to determine the effectiveness, reliability, and adaptability of the

cascaded model in real-world safety monitoring scenarios. The evaluation focused on the model's detection performance and both Level-I and Level-II classifications. To measure the performance, we used several key metrics: mean average precision (mAP) at different intersection over union (IoU) thresholds, precision, recall, accuracy, and F1 score, as outlined in Equations (5-8) – (5-11).

The IoU metric helps to assess how well the predicted bounding boxes align with the actual ones. A perfect match resulted in an IoU of one. We used average precision (AP) to measure detection precision at various thresholds, whereas mAP provided an overall measure across all categories and IoU thresholds. To evaluate the model's classification performance, we analyzed the confusion matrix to calculate precision, recall, F1-score, and overall accuracy.

This comprehensive set of metrics allowed us to thoroughly evaluate the capabilities of our cascaded model in construction site environments. Analyzing these indicators gave us valuable insights into the effectiveness of the model in enhancing safety monitoring, advancing SSSS technologies, and on-site safety protocols. This approach also helped us identify the model's strengths and potential areas for improvement when applied to real-world situations.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \tag{5-8}$$

$$AP = \sum_n (R_n - R_{n-1}) P_n \tag{5-9}$$

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \tag{5-10}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5-11}$$

$$F1\ Score = 2\left(\frac{Precision * Recall}{Precision + Recall}\right) \tag{5-12}$$

In these equations: R_n is the recall at the *nth* threshold, R_{n-1} is the recall at the previous threshold, P_n is the precision at the *nth* threshold. Other variables are defined in Equations (5-1) – (5-6).

Table 5-5 provides a detailed summary of the performance metrics for the cascaded model in object detection, as well as the Level-I and Level-II hierarchical classifications. The metrics are presented for the mobile and tower crane classes and are also averaged across all classes. For object detection, four key metrics were assessed: precision, recall, mAP50, and mAP50-95. The mAP50 metric represents the mean average precision at a 50% Intersection over Union (IoU), indicating the precision of the model when detections overlap with ground-truth bounding boxes by at least 50%. The mAP50-95 metric calculates the mean average precision across multiple IoU thresholds ranging from 50% to 95% in 5% increments. These metrics evaluate the detection abilities of the model and their accuracy in defining crane-type boundaries across various IoU thresholds. The cascaded model demonstrated an overall precision of 92.10% for crane detection, 90.90% for mobile cranes, and 93.30% for tower cranes. In addition, the model achieved 93.80% recall, 96.80% mAP50, and 96.80% mAP50-95. Four metrics were considered for Level-I and Level-II classifications: precision, recall, accuracy, and F1 score. In the level-I classifications, the overall precision, recall, accuracy, and F1 scores were 99.25%, 99.25%, 99.25%, and 98.50%, respectively. In the level-II classifications, the overall precision, recall, accuracy, and F1 scores were 99.58%, 99.42%, 99.47%, and 99.47%, respectively.

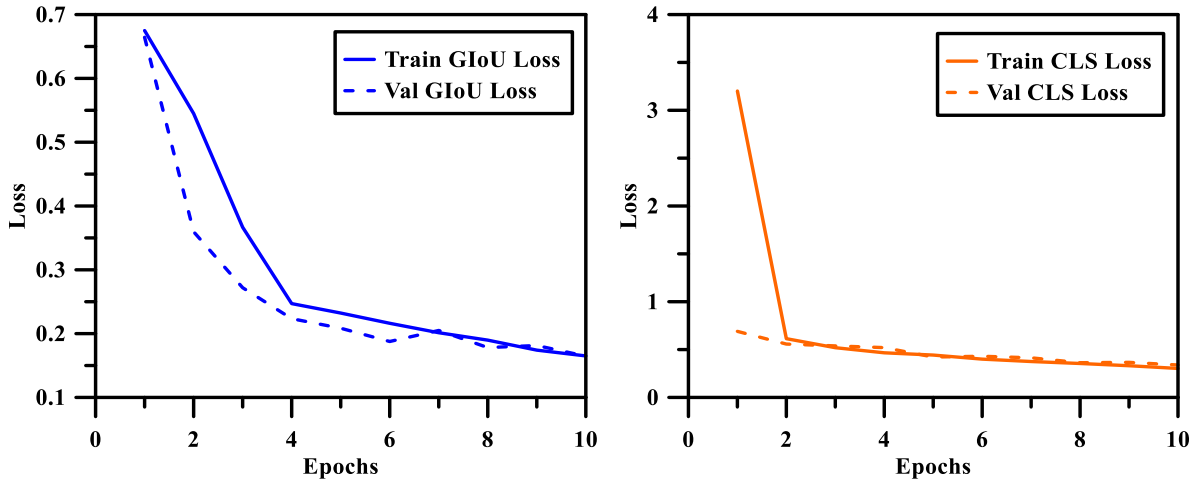
Higher IoU values indicate close alignment between the predicted and actual object boundaries. The AP and mAP values suggest consistent precision across the different crane types and confidence thresholds. Higher IoU thresholds, such as 75% or above, reflect strict requirements for overlap, demonstrating high precision in object localization. This is critical in construction settings where the accurate detection of equipment and personnel is vital. Lower IoU

thresholds, such as 50%, still offer valuable insights, particularly in complex environments where achieving perfect overlap may be difficult due to occlusion or rapid movement.

Table 5. 5: Performance metrics of the cascaded model.

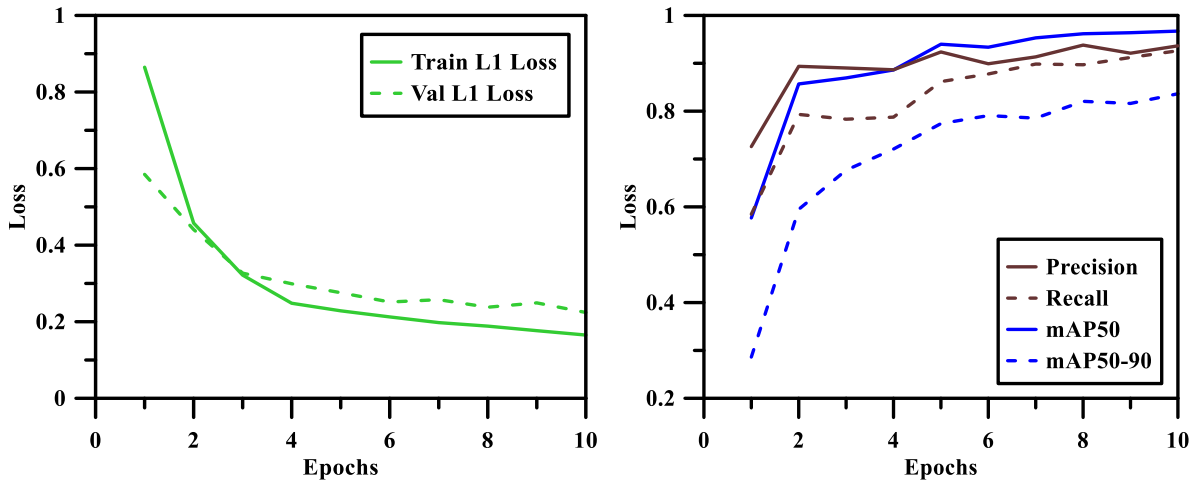
Objective Operations	Types/Activities	Performance metrics					
		Precision	Recall	mAP50	mAP50-95	Accuracy	F1 Score
Crane detection	Overall	92.10	93.80	96.80	96.80	--	--
	Mobile crane	90.90	93.10	96.50	96.50	--	--
	Tower crane	93.30	94.50	97.10	97.10	--	--
Level-I classification	Overall	99.25	99.25	--	--	99.25	98.50
	Safe operations	99.00	99.00	--	--	99.00	99.00
	Unsafe operations	99.00	99.00	--	--	99.00	96.00
	Near-miss incidents	100.00	100.00	--	--	100.00	100.00
	Incidents	99.00	99.00	--	--	99.00	99.00
Level-II classification	Overall	99.58	99.42	--	--	99.47	99.47
	Class_0(b_{41})	100.00	100.00	--	--	100.00	100.00
	Class_1(b_{42})	98.00	92.00	--	--	95.00	95.00
	Class_2(b_{43})	100.00	100.00	--	--	100.00	100.00
	Class_3(b_{44})	96.00	99.00	--	--	97.00	97.00
	Class_4(b_{45})	99.00	100.00	--	--	99.00	99.00
	Class_5(b_{46})	100.00	100.00	--	--	100.00	100.00
	Class_6(b_{47})	100.00	100.00	--	--	100.00	100.00
	Class_7(b_{31})	100.00	100.00	--	--	100.00	100.00
	Class_8(b_{32})	100.00	100.00	--	--	100.00	100.00
	Class_9(b_{33})	100.00	100.00	--	--	100.00	100.00
	Class_10(b_{11})	100.00	100.00	--	--	100.00	100.00
	Class_11(b_{12})	99.00	100.00	--	--	99.00	99.00
	Class_12(b_{13})	100.00	99.00	--	--	100.00	100.00
	Class_13(b_{14})	100.00	100.00	--	--	100.00	100.00
	Class_14(b_{21})	100.00	100.00	--	--	100.00	100.00
	Class_15(b_{22})	100.00	100.00	--	--	100.00	100.00
	Class_16(b_{23})	100.00	100.00	--	--	100.00	100.00
	Class_17(b_{24})	100.00	100.00	--	--	100.00	100.00
	Class_18(b_{25})	100.00	99.00	--	--	100.00	100.00

Figure 5-12 illustrates the various loss functions and evaluation metrics that are crucial for training and assessing object detection models. These include Generalized Intersection over Union (GIoU), classification loss (CLS), L1 (Smooth L1) loss, precision, recall, mAP50, and mAP50-95, as detailed in Section 5.3.4.1. These functions play vital roles in training the model, particularly for complex tasks such as identifying objects in images and determining their locations. Each function explicitly guides the learning process of the model. GIoU improves the accuracy of bounding box predictions. It provides a more informative and smoother gradient during training, which helps the model to make better predictions for object locations. CLS loss penalizes incorrect predictions of object classes within the detected bounding boxes. It employs cross-entropy to measure the difference between the predicted class probabilities and actual class labels. This loss helps the model detect objects and recognize specific types, such as mobile or tower cranes. L1 loss indicates the accuracy of positive predictions. Precision, recall, mAP50, and mAP50-95 evaluate the ability of the model to accurately identify, localize, and classify objects across different thresholds and scenarios. By incorporating these loss functions and evaluation metrics, the model can be effectively trained and assessed for its performance in object detection tasks.



(a) GIoU loss function

(b) CLS loss function



(c) L1 loss function

(d) Precision, recall, mAP50, and mAP50-95

Figure 5. 12: Loss functions.

5.6 Discussions

This section compares the proposed cascaded model against other state-of-the-art detection and classification models for risk identification. The evaluation covered various aspects, including computational complexity, accuracy, precision, reliability, and adaptability. By assessing these critical metrics, we aimed to provide a comprehensive assessment of the effectiveness of our model in relation to existing approaches in the field.

5.6.1 Object detection

The proposed cascaded model was compared with other leading object detection models under consistent training and testing conditions to provide a fair evaluation. This included using the same number of epochs, experimental setup, and datasets for training and validation. Table 5-6 details the performance comparison between our cascaded transformer-based model and notable models, such as YOLOv11, YOLOv10, YOLOv9, YOLOv8, YOLO-NAS, and DETR. The results indicate that our cascaded model (RT-DETR-L) outperforms the other models in several metrics, achieving a precision of 92.10%, recall of 93.80%, mAP50 of 96.80%, and mAP50-95 of 96.80%. For instance, compared with YOLOv9, which achieved a precision of

87.60%, recall of 75.90%, mAP50 of 84.50%, and mAP50-95 of 67.50%, our cascaded model showed improvements of 4.50% in precision, 17.90% in recall, 12.30% in mAP50, and 16.90% in mAP50-95. These results are inspiring, given the model's performance on challenging low-quality CCTV footage typical of construction sites. Although other state-of-the-art models may perform adequately under standard conditions, they often face difficulties with object detection in low-resolution images.

In contrast, the cascaded model maintains high accuracy across both standard and low-resolution images. It also performs well when processing images from complex and dynamic construction environments, with minimal performance degradation. This robust capability across varying image qualities and environmental conditions underscores the adaptability and effectiveness of the proposed model in real-world construction site scenarios.

Table 5. 6: Object detection comparison between the cascaded model and other models.

Model	Precision	Recall	mAP50	mAP50-95
YOLOv11	85.50	74.10	83.30	66.20
YOLOv10	77.80	75.60	83.90	68.00
YOLOv9	87.60	75.90	84.50	67.50
YOLOv8	86.90	69.60	81.60	63.70
YOLO-NAS	79.95	64.03	75.07	58.60
DETR	86.29	74.76	83.23	66.49
Our model (RT-DETR-L)	92.10	93.80	96.80	96.80

5.6.2 Hierarchical classification

Table 5-7 compares a cascaded transformer-based model against other leading models, such as ResNet34, DiNAT, YOLO-NAS, and DETR, specifically to classify crane operations into safe, unsafe, near-miss, and incidents categories. The evaluation included several key performance metrics: precision, recall, accuracy, and F1 score. The cascaded model demonstrated superior performance, achieving a precision of 99.25%, recall of 99.25%, accuracy of 99.25%, and F1

scores of 98.50%. For instance, compared to DiNAT, which also performed well, the precision was 95.50%, recall was 95.30%, accuracy was 95.60%, and F1 score was 95.60%. The cascaded model surpassed DiNAT by 3.75% in precision, 3.95% in recall, 3.65% in accuracy, and 2.90% in F1 score. This exceptional performance is attributed to the unique architecture of the proposed model, which supports advanced hierarchical classification capabilities even in low-quality images. The results highlight the effectiveness of the proposed model in accurately distinguishing between safe, unsafe, near-miss, and incident crane operations, surpassing other state-of-the-art methods in this crucial classification task.

Table 5. 7: Level-I safety status classification comparison.

Model	Precision	Recall	Accuracy	F1 Score
Resnet34	92.80	92.90	91.70	91.70
DiNAT	95.50	95.30	95.60	95.60
YOLO-NAS	93.60	93.70	92.40	92.40
DETR	93.10	93.20	91.90	91.90
Our model (DINOv2)	99.25	99.25	99.25	98.50

Table 5-8 provides a comparison of the performance of the cascaded model with other advanced models for classifying specific crane operation activities. The cascaded model achieved notable results, including a precision of 99.58%, recall of 99.42%, accuracy of 99.47%, and an F1 score of 99.47%. These results outperform the other models in all metrics. The effectiveness of the cascaded model in differentiating between various specific activities is attributed to its multi-layered approach. The ViT component, tailored for detailed action classification, complements the accurate object detection capabilities of RT-DETR-L. This combination and the hierarchical classification strategy facilitated the precise classification of activity types.

Table 5. 8: Level-II specific activity classification comparison.

Model	Precision	Recall	Accuracy	F1 Score
Resnet34	95.60	97.75	97.60	97.60

DiNAT	95.40	95.04	94.70	94.70
YOLO-NAS	92.40	91.91	93.22	93.22
DETR	99.18	97.06	97.96	97.96
Our model (ViT)	99.58	99.42	99.47	99.47

5.6.3 Computational complexity

The analysis focused on the individual models and their integration within the complete pipeline to evaluate the computational complexity of the cascaded transformer-based model. This evaluation examines the distinct computational and memory requirements of each model, which collectively affect the efficiency of the pipeline designed for detecting and classifying unsafe crane operations in construction site environments. The assessment utilized various metrics, including inference time, inference speed, resident set size (RSS), virtual memory size (VMS), and Giga floating-point operations (GFLOPs).

The inference time represents the duration required to process a single instance through the model, which is critical for applications that demand rapid data processing. This provided insights into the responsiveness and speed of the computational framework. On the other hand, inference speed indicates the number of outputs a model can produce per unit of time. Memory utilization was gauged using the RSS and VMS, where RSS measures the amount of memory a process occupies in RAM, excluding any swapped-out memory, and VMS accounts for the total virtual memory allocated by the process, including both used physical memory and swapped-out space. These metrics are vital for evaluating the scalability and memory efficiency of the models, particularly when deployed in systems with constrained memory resources. GFLOPs measure computational complexity by quantifying the billion floating-point operations performed per second, directly assessing the processing power required and enabling comparison of different components and configurations.

Table 5-9 presents the computational characteristics of each model within the cascaded pipeline. The SRGAN, tasked with resolution enhancement, exhibits the highest computational demands, with a lengthy average inference time of 0.40 seconds and a high GFLOPs count of 686.57, reflecting its resource-intensive nature. In contrast, DINOv2, despite its advanced multiscale processing capabilities, has a faster inference time of 0.06 seconds and consumes less memory, as indicated by its RSS of 3.33 MB and VMS of 4.27 MB. RT-DETR-L and ViT are notably efficient, with inference times of 0.13 seconds and 0.11 seconds, respectively. They demonstrate lower resource utilization with minimal RSS and VMS and reduced GFLOPs, suggesting they perform effectively even with limited computational resources while maintaining rapid processing speeds. The integration of these models into a unified pipeline results in a total inference time of 0.70 seconds, an inference speed of 1.43 seconds, an RSS of 439.29 MB, a VMS of 598.75 MB, and a high GFLOPs of 1000.41. The memory usage of the cascaded pipeline, as shown by the RSS and VMS, indicates substantial system resource requirements. However, this resource intensity is offset by the advanced capabilities of the pipeline and its high performance in practical applications.

In conclusion, the cascaded model is a promising solution for detecting and classifying unsafe crane operations in dynamic construction site environments. This integrated model significantly improves crane safety standards by combining enhanced resolution, accurate object detection, and detailed hierarchical classification.

Table 5. 9: Computational time comparison between the cascaded model and other models.

Model	Inference time (s)	Inference speed (s)	RSS (MB)	VMS (MB)	GFLOPs
SRGAN	0.40	2.49	187.94	202.40	686.57
RT-DETR-L	0.13	7.69	244.20	388.51	264.78
DINOv2	0.06	16.67	3.33	4.27	28.04
ViT	0.11	9.09	3.82	3.57	21.02
Cascaded model	0.70	1.43	439.29	598.75	1000.41

5.7 Chapter Summary

This chapter introduces a novel cascaded learning approach for automated safety risk assessment in crane operations within dynamic construction environments. It addresses the inadequacies of conventional safety monitoring, which relies on error-prone and delayed manual reporting. Integrating advanced computer vision techniques, the model combines four components: Super-Resolution Generative Adversarial Network (SRGAN) for image enhancement, Real-Time Detection Transformer-Large (RT-DETR-L) for crane detection, DINOv2 for safety classification, and Vision Transformer (ViT) for detailed activity classification.

The cascaded framework effectively processes low-resolution surveillance footage to deliver high-precision safety insights. It identifies crane operations as safe, unsafe, near-miss incidents, or incidents, and categorizes specific activities. Safety risks are assessed through confidence levels, probability matrices, and severity ratings, with real-time warnings issued for high-risk activities. The system operates with detection precision of 92.10%, safety status classification accuracy of 99.25%, and specific activity classification accuracy of 99.47%.

A robust dataset of crane operations, including diverse scenarios of safe and unsafe activities, underpins the model's training. Experimental results highlight significant improvements in image quality (PSNR: 15.49 dB, SSIM: 0.87) and detection metrics (mAP50: 96.80%). The cascaded model's performance surpasses existing methods in computational efficiency, accuracy, and adaptability, achieving inference times of 0.70 seconds.

This study advances Smart Site Safety Systems (4S) by providing a comprehensive, scalable solution for real-time hazard detection and mitigation. Its contributions lie in transforming conventional safety protocols into proactive systems, optimizing construction safety management, and laying the groundwork for integrating artificial intelligence in construction

risk assessment. Future research will focus on extending this methodology to other construction equipment and scenarios, addressing site-specific challenges, and enhancing system scalability.

CHAPTER 06

CONCLUSIONS, LIMITATIONS, AND FUTURE WORKS

6.1 Conclusions

This thesis presents a comprehensive investigation into advanced solutions for enhancing the safety, efficiency, and reliability of crane operations in construction environments. First, it introduces an integrated computing system for occlusion-free and collision-free crane lift path planning and replanning in dynamic environments. The system leverages a hybrid algorithm combining A* and genetic algorithms (GA) alongside raycasting techniques to ensure efficient pathfinding and real-time decision-making. The methodology significantly improves path smoothness, computational efficiency, and scalability through asynchronous communication, enhanced heuristic information, and optimized subproblem division. The physics-based simulation platform further enhances practical applicability by providing 4D crane-lift animations, aiding construction planners and operators in optimizing lift paths for safe project delivery.

Second, the study addresses the risks associated with structural fatigue and aging-induced deterioration in construction tower cranes through a digital twin-driven (DTD) framework. This approach successfully predicts the degraded lifting capacities of aging cranes using real-time load and deflection data, ensuring safety during lifting operations. The DTD framework demonstrated exceptional accuracy, with MSE and R^2 values of 0.2253 and 0.9973, respectively, in predicting lifting capacities for scaled-down prototypes and actual cranes. These findings highlight the potential of DTD approaches in improving safety, efficiency, and real-time decision-making for aging cranes, significantly enhancing workplace monitoring and operational safety.

Lastly, the thesis proposes a cascaded transformer-based model for automated safety risk assessment during crane operations. The model integrates advanced neural networks—SRGAN, RT-DETR-L, DINOv2, and ViT—for high-resolution image processing, crane detection, safety status classification, and activity recognition. The model achieves notable performance, with detection precision of 92.10% and classification accuracies of 99.25% and 99.47% for safety statuses and activities, respectively. This robust system enhances real-time safety monitoring and decision-making, advancing the Smart Site Safety System (SSSS) and contributing significantly to safety management in dynamic construction settings.

6.2 Limitations

Despite the promising contributions of this thesis, several limitations exist that warrant attention for future development. These limitations span across the three core components presented: the integrated computing system, the digital twin-driven framework, and the cascaded transformer-based safety risk assessment model.

The system demonstrates a high level of automation and computational efficiency through the hybrid A*-GA algorithm. However, its practical implementation remains challenged by system complexity. Designing a fully integrated solution that seamlessly connects sensing, decision-making, and actuation modules requires substantial engineering effort. Currently, the system partially depends on manual intervention, especially for initializing crane and obstacle parameters. Furthermore, while the hybrid algorithm enhances planning performance, its scalability and adaptability to various crane types (e.g., crawler, gantry, or mobile cranes) and cross-industry applications (e.g., shipping, manufacturing, or mining) have not been fully explored or validated in real-world deployments. The absence of real-time feedback loops and adaptive behaviors further limits its responsiveness in rapidly changing site environments.

This framework effectively simulates the degradation trends of aging tower cranes. However, it is constrained by reliance on scaled-down physical prototypes, which may not capture the mechanical and structural behavior of full-scale cranes. Key components such as the crane's foundation, slewing unit, and climbing frame are excluded from current models, potentially affecting the overall structural integrity and predictive accuracy. The integration between virtual and physical systems remains semi-automated, requiring human effort to synchronize data and states. Moreover, the lack of incorporation of real-world operational loads, environmental stresses (e.g., wind, temperature), and usage histories into the digital twin limits the generalizability of the simulation outcomes.

While the model performs well in controlled environments, its performance in complex, real-world construction sites remains uncertain. The current framework relies on manually selected frames for analysis, which introduces inefficiencies and potential biases. Additionally, the model has been trained and tested on a limited range of equipment types and working conditions, reducing its robustness across diverse construction scenarios. The system also lacks integration with multimodal data sources—such as audio, thermal imaging, or vibration data—which could enhance context-awareness and improve detection capabilities. Furthermore, the absence of a real-time deployment mechanism limits the potential for proactive risk mitigation on-site.

6.3 Future Works

In response to the identified limitations, it is recommended that future research explore the following directions to further enhance the robustness, generalizability, and practical relevance of the developed methodologies:

It is recommended that future research explore reducing the complexity of system integration by developing modular, plug-and-play architectures with standardized communication

protocols. Full automation should be pursued, including automatic detection and labeling of obstacles and cranes from sensor inputs (e.g., LiDAR, RGB-D cameras). The inclusion of project-specific constraints, such as site layout, lifting plans, and weather data, will improve context sensitivity. Expanding the system's applicability to different types of cranes and construction scenarios across various industries is essential for broader utility. Additionally, incorporating adaptive learning algorithms and reinforcement learning techniques could allow the system to dynamically respond to unanticipated environmental changes and optimize paths in real time.

It is recommended that future research explore the development of full-scale digital twins by incorporating comprehensive structural elements such as the crane base, slewing unit, climbing frame, and counterweights to improve the realism of simulations. Automated bidirectional communication between virtual and physical components using IoT protocols and cloud-based platforms should be established to facilitate real-time monitoring and control. The integration of real operational data (e.g., stress-strain measurements, fatigue cycles, and maintenance logs) into predictive models will increase accuracy. Coupling these models with advanced finite element simulations and machine learning algorithms may allow early warning of performance degradation, thereby enabling condition-based maintenance strategies.

It is recommended that future research explore automating the selection and preprocessing of input data using object detection and scene segmentation techniques to minimize manual effort. The model should be expanded to handle a wider range of construction equipment (e.g., excavators, hoists, forklifts) and more complex environmental contexts (e.g., nighttime operations, foggy weather). Integrating multimodal data streams—including audio cues (alarms, worker communication), motion sensors, GPS data, and environmental sensors—can significantly enhance situational awareness. Real-time deployment using edge-computing

frameworks and closed-loop feedback mechanisms would support immediate hazard detection and response. Finally, deploying this system in live construction projects and evaluating its effectiveness in real-world operations would provide valuable insights into its practical adoption and impact on safety performance.

REFERENCES

Abdelmageed, S., & Zayed, T. (2020). A study of literature in modular integrated construction - Critical review and future directions. In Journal of Cleaner Production (Vol. 277). Elsevier Ltd. doi: 10.1016/j.jclepro.2020.124044

Ajeil, F. H., Ibraheem, I. K., Azar, A. T., & Humaidi, A. J. (2020). Grid-based mobile robot path planning using aging-based ant colony optimization algorithm in static and dynamic environments. *Sensors (Switzerland)*, *20*(7). doi: 10.3390/s20071880

Ajmal, M. S., Ali, D., Ramesh Babu, ; N, & Varghese, K. (n.d.). *Collision Free Path Planning of Cooperative Crane Manipulators Using Genetic Algorithm*. doi: 10.1061/ASCE0887-3801200519:2182

AlBahnassi, H., & Hammad, A. (2012). Near Real-Time Motion Planning and Simulation of Cranes in Construction: Framework and System Architecture. *Journal of Computing in Civil Engineering*, *26*(1), 54–63. doi: 10.1061/(asce)cp.1943-5487.0000123

Al-Faris, M., Chiverton, J., Ndzi, D., & Ahmed, A. I. (2020). A review on computer vision-based methods for human action recognition. In Journal of Imaging (Vol. 6, Issue 6). MDPI. doi: 10.3390/jimaging6060046

Al-Hussein, M., Athar Niaz, M., Yu, H., & Kim, H. (2006). Integrating 3D visualization and simulation for tower crane operations on construction sites. *Automation in Construction*, *15*(5), 554–562. doi: 10.1016/j.autcon.2005.07.007

Ali, A. H., Zayed, T., & Hussein, M. (2024). Crane safety operations in modular integrated construction. In Automation in Construction (Vol. 164). Elsevier B.V. doi: 10.1016/j.autcon.2024.105456

Ali, A. H., Zayed, T., Wang, R. D., & Kit, M. Y. S. (2024). Tower crane safety technologies: A synthesis of academic research and industry insights. In Automation in Construction (Vol. 163). Elsevier B.V. doi: 10.1016/j.autcon.2024.105429

Ali, M. S. A. D., Babu, N. R., & Varghese, K. (2005). Collision Free Path Planning of Cooperative Crane Manipulators Using Genetic Algorithm. *Journal of Computing in Civil Engineering*, *19*(2), 182–193. doi: 10.1061/(asce)0887-3801(2005)19:2(182)

Alsakka, F., Assaf, S., El-Chami, I., & Al-Hussein, M. (2023). Computer vision applications in offsite construction. In Automation in Construction (Vol. 154). Elsevier B.V. doi: 10.1016/j.autcon.2023.104980

Anwar, G. A., Akber, M. Z., Ahmed, H. A., Hussain, M., Nawaz, M., Anwar, J., Chan, W.-K., & Lee, H.-H. (2024). Life-Cycle Performance Modeling for Sustainable and Resilient Structures under Structural Degradation: A Systematic Review. *Buildings*, *14*(10), 3053. doi: 10.3390/buildings14103053

Anwar, G. A., Hussain, M., Akber, M. Z., Khan, M. A., & Khan, A. A. (2023). Sustainability-Oriented Optimization and Decision Making of Community Buildings under Seismic Hazard. *Sustainability (Switzerland)*, *15*(5). doi: 10.3390/su15054385

Arumugaprabu, V., Ajith, S., Jerendran, J., Naresh, K., & Rama Sreekanth, P. S. (2022). Hazard identification and risk assessment using integrated exposure frequency and legislation requirements (HIRA-FL) in construction sites. *Materials Today: Proceedings*, *56*, 1247–1250. doi: 10.1016/j.matpr.2021.11.178

Aslan, M. F., Durdu, A., & Sabanci, K. (2023). Goal distance-based UAV path planning approach, path optimization and learning-based path estimation: GDRRT*, PSO-GDRRT* and BiLSTM-PSO-GDRRT*. *Applied Soft Computing*, *137*. doi: 10.1016/j.asoc.2023.110156

Awolusi, I., Marks, E., & Hallowell, M. (2018). Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices. *Automation in Construction*, *85*, 96–106. doi: 10.1016/j.autcon.2017.10.010

Bandara, C. S., Siriwardane, S. C., Dissanayake, U. I., & Dissanayake, R. (2015). Developing a full range S-N curve and estimating cumulative fatigue damage of steel elements. *Computational Materials Science*, *96*(PA), 96–101. doi: 10.1016/j.commatsci.2014.09.009

Bansal, S., Chen, M., Fisac, J. F., & Tomlin, C. J. (2016). *Safe Sequential Path Planning Under Disturbances and Imperfect Information*. Retrieved from http://arxiv.org/abs/1603.05208

Benkabouche, S., Guechichi, H., Amrouche, A., & Benkhettab, M. (2015). A modified nonlinear fatigue damage accumulation model under multiaxial variable amplitude loading. *International Journal of Mechanical Sciences*, *100*, 180–194. doi: 10.1016/j.ijmecsci.2015.06.016

Bertram, D., Kuffner, J., Dillmann, R., & Asfour, T. (2006a). *An Integrated Approach to Inverse Kinematics and Path Planning for Redundant Manipulators*.

Bertram, D., Kuffner, J., Dillmann, R., & Asfour, T. (2006b). *An Integrated Approach to Inverse Kinematics and Path Planning for Redundant Manipulators*.

Betit, E., Barlet, G., & Bunting, J. (2022). *Struck-by Hazards, Barriers, and Opportunities in the Construction Industry CPWR - The Center for Construction Research and Training*. *https://www.cpwr.com*.

Brown, S., Brooks, R. D., & Sue Dong, X. (2020). *Nonfatal Injury Trends in the Construction Industry-CPWR - The Center for Construction Research and Training*. *https://www.cpwr.com*.

Brown, S., Brown, S., & Brown, S. (2021). *Fatal Injury Trends in the Construction Industry. CPWR - The Center for Construction Research and Training. https://www.cpwr.com*.

Brozovsky, J., Labonnote, N., & Vigren, O. (2024). Digital technologies in architecture, engineering, and construction. In Automation in Construction (Vol. 158). Elsevier B.V. doi: 10.1016/j.autcon.2023.105212

Bruce, J., & Veloso, M. (n.d.). *Real-Time Randomized Path Planning for Robot Navigation* *.

Bureau of Labor Statistics. (2023). *NEWS RELEASE - NATIONAL CENSUS OF FATAL OCCUPATIONAL INJURIES IN 2022 - US DEPARTMENT OF LABOR*. Retrieved from www.bls.gov/iif

Burg, L., Lino, C., Christie, M., & Christophe Lino, F. (2021). *Real-Time Cinematic Tracking of Targets in Dynamic Environments*. Retrieved from https://inria.hal.science/hal-03519261

Cai, P., Cai, Y., Chandrasekaran, I., & Zheng, J. (2016a). Parallel genetic algorithm based automatic path planning for crane lifting in complex environments. In Automation in Construction (Vol. 62, pp. 133–147). Elsevier. doi: 10.1016/j.autcon.2015.09.007

Cai, P., Cai, Y., Chandrasekaran, I., & Zheng, J. (2016b). Parallel genetic algorithm based automatic path planning for crane lifting in complex environments. In Automation in Construction (Vol. 62, pp. 133–147). Elsevier. doi: 10.1016/j.autcon.2015.09.007

Cai, P., Chandrasekaran, I., Zheng, J., & Cai, Y. (2018). Automatic Path Planning for Dual-Crane Lifting in Complex Environments Using a Prioritized Multiobjective PGA. *IEEE Transactions on Industrial Informatics*, *14*(3), 829–845. doi: 10.1109/TII.2017.2715835

Cantrell, J. H. (2006). Dependence of microelastic-plastic nonlinearity of martensitic stainless steel on fatigue damage accumulation. *Journal of Applied Physics*, *100*(6). doi: 10.1063/1.2345614

Cao, X., Zhao, G., & Xiao, W. (2022). Digital Twin–oriented real-time cutting simulation for intelligent computer numerical control machining. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *236*(1–2), 5–15. doi: 10.1177/0954405420937869

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). *End-to-End Object Detection with Transformers*. Retrieved from http://arxiv.org/abs/2005.12872

Cekmez, U., Ozsiginan, M., & Sahingoz, O. K. (2014). A UAV path planning with parallel ACO algorithm on CUDA platform. *2014 International Conference on Unmanned*

Aircraft Systems, ICUAS 2014 - Conference Proceedings, 347–354. doi: 10.1109/ICUAS.2014.6842273

Chaboche, J. L., & Lesne, P. M. (1988). A NON-LINEAR CONTINUOUS FATIGUE DAMAGE MODEL. In Fatigue From Engng Muter. Strucf (Issue 1).

Chang, Y. C., Hung, W. H., & Kang, S. C. (2012a). A fast path planning method for single and dual crane erections. *Automation in Construction*, *22*, 468–480. doi: 10.1016/j.autcon.2011.11.006

Chang, Y. C., Hung, W. H., & Kang, S. C. (2012b). A fast path planning method for single and dual crane erections. *Automation in Construction*, *22*, 468–480. doi: 10.1016/j.autcon.2011.11.006

Chen, A., Jacob, M., Shoshani, G., & Charter, M. (2023). Using computer vision, image analysis and UAVs for the automatic recognition and counting of common cranes (Grus grus). *Journal of Environmental Management*, *328*. doi: 10.1016/j.jenvman.2022.116948

Chen, D. L., Giordano, M., Benko, H., Grossman, T., & Santosa, S. (2023, October 9). GazeRayCursor: Facilitating Virtual Reality Target Selection by Blending Gaze and Controller Raycasting. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*. doi: 10.1145/3611659.3615693

Chen, M., Fisac, J. F., Sastry, S., & Tomlin, C. J. (2014). *Safe Sequential Path Planning of Multi-Vehicle Systems via Double-Obstacle Hamilton-Jacobi-Isaacs Variational Inequality*. Retrieved from http://arxiv.org/abs/1412.7223

Chen, W., Li, C., & Guo, H. (2023). A lightweight face-assisted object detection model for welding helmet use. *Expert Systems with Applications*, *221*. doi: 10.1016/j.eswa.2023.119764

Chen, Y., Cutler, M., & How, J. P. (2015). Decoupled multiagent path planning via incremental sequential convex programming. *Proceedings - IEEE International Conference on Robotics and Automation*, *2015-June*(June), 5954–5961. doi: 10.1109/ICRA.2015.7140034

Cheng, D. J., Zhang, J., Hu, Z. T., Xu, S. H., & Fang, X. F. (2020). A Digital Twin-Driven Approach for On-line Controlling Quality of Marine Diesel Engine Critical Parts. *International Journal of Precision Engineering and Manufacturing*, *21*(10), 1821–1841. doi: 10.1007/s12541-020-00403-y

Chian, E. Y. T., Goh, Y. M., Tian, J., & Guo, B. H. W. (2022). Dynamic identification of crane load fall zone: A computer vision approach. *Safety Science*, *156*. doi: 10.1016/j.ssci.2022.105904

Chiang, Y.-H., Francis, ;, Wong, K.-W., & Liang, S. (2017). *Fatal Construction Accidents in Hong Kong*. doi: 10.1061/(ASCE)CO.1943

Christiand, G. K. (2020). Digital twin approach for tool wear monitoring of micro-milling. *Procedia CIRP*, *93*, 1532–1537. doi: 10.1016/j.procir.2020.03.140

Code of Practice for Safe Use of Tower Cranes. (n.d.). Retrieved from http://www.labour.gov.hk/eng/public/content2_8b.htm.http://www.labour.gov.hk/eng/tele/osh.htmor

Dakulovi, M., & Petrovi, I. (2011). Two-way D* algorithm for path planning and replanning. *Robotics and Autonomous Systems*, *59*(5), 329–342. doi: 10.1016/j.robot.2011.02.007

Das, S., Mukhopadhyay, G., & Bhattacharyya, S. (2018). Failure analysis of a 40 ton crane hook at a Hot Strip Mill. *MATEC Web of Conferences*, *165*. doi: 10.1051/matecconf/201816510006

Deebak, B. D., & Al-Turjman, F. (2021). Digital-twin assisted: Fault diagnosis using deep transfer learning for machining tool condition. *International Journal of Intelligent Systems*. doi: 10.1002/int.22493

Dobrucali, E., Demirkesen, S., Sadikoglu, E., Zhang, C., & Damci, A. (2024). Investigating the impact of emerging technologies on construction safety performance. *Engineering, Construction and Architectural Management*, *31*(3), 1322–1347. doi: 10.1108/ECAM-07-2022-0668

Dong, Q., He, B., Qi, Q., & Xu, G. (2021a). Real-time prediction method of fatigue life of bridge crane structure based on digital twin. *Fatigue and Fracture of Engineering Materials and Structures*, *44*(9), 2280–2306. doi: 10.1111/ffe.13489

Dong, Q., He, B., Qi, Q., & Xu, G. (2021b). Real-time prediction method of fatigue life of bridge crane structure based on digital twin. *Fatigue and Fracture of Engineering Materials and Structures*, *44*(9), 2280–2306. doi: 10.1111/ffe.13489

Dong, R., Hu, M., Cui, T., Wan, D., & Meng, X. (2023). *A Mathematical Modeling Approach for Optimal Parking Space Selection and Path Planning in Autonomous Parking Systems with UAV-Assisted Topsis Entropy Weight Method*. doi: 10.21203/rs.3.rs-3639234/v1

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. Retrieved from http://arxiv.org/abs/2010.11929

Duan, R., Deng, H., Tian, M., Deng, Y., & Lin, J. (2022). SODA: A large-scale open site object detection dataset for deep learning in construction. *Automation in Construction*, *142*. doi: 10.1016/j.autcon.2022.104499

Duarte, J., Torres Marques, A., & Santos Baptista, J. (2021). Occupational accidents related to heavy machinery: A systematic review. *Safety*, *7*(1). doi: 10.3390/safety7010021

Dutta, S., Cai, Y., Huang, L., & Zheng, J. (2020a). Automatic re-planning of lifting paths for robotized tower cranes in dynamic BIM environments. *Automation in Construction*, *110*. doi: 10.1016/j.autcon.2019.102998

Dutta, S., Cai, Y., Huang, L., & Zheng, J. (2020b). Automatic re-planning of lifting paths for robotized tower cranes in dynamic BIM environments. *Automation in Construction*, *110*. doi: 10.1016/j.autcon.2019.102998

Elgendi, E.-B. O., Shawki, K. M., & Ashraf Mohy, A. (2023). Video analysis for tower crane production rate estimation. *Journal of Information Technology in Construction*, *28*, 138–150. doi: 10.36680/j.itcon.2023.007

Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., Rose, T. M., & An, W. (2018). Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Automation in Construction*, *85*, 1–9. doi: 10.1016/j.autcon.2017.09.018

Fang, W., Ding, L., Love, P. E. D., Luo, H., Li, H., Peña-Mora, F., Zhong, B., & Zhou, C. (2020a). Computer vision applications in construction safety assurance. In Automation in Construction (Vol. 110). Elsevier B.V. doi: 10.1016/j.autcon.2019.103013

Fang, W., Ding, L., Love, P. E. D., Luo, H., Li, H., Peña-Mora, F., Zhong, B., & Zhou, C. (2020b). Computer vision applications in construction safety assurance. In Automation in Construction (Vol. 110). Elsevier B.V. doi: 10.1016/j.autcon.2019.103013

Fang, W., Ding, L., Zhong, B., Love, P. E. D., & Luo, H. (2018). Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach. *Advanced Engineering Informatics*, *37*, 139–149. doi: 10.1016/j.aei.2018.05.003

Fang, W., Love, P. E. D., Ding, L., Xu, S., Kong, T., & Li, H. (2023). Computer Vision and Deep Learning to Manage Safety in Construction: Matching Images of Unsafe Behavior and Semantic Rules. *IEEE Transactions on Engineering Management*, *70*(12), 4120–4132. doi: 10.1109/TEM.2021.3093166

Fang, W., Love, P. E. D., Luo, H., & Ding, L. (2020). Computer vision for behaviour-based safety in construction: A review and future directions. In Advanced Engineering Informatics (Vol. 43). Elsevier Ltd. doi: 10.1016/j.aei.2019.100980

Fang, W., Zhong, B., Zhao, N., Love, P. E. D., Luo, H., Xue, J., & Xu, S. (2019). A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network. *Advanced Engineering Informatics*, *39*, 170–177. doi: 10.1016/j.aei.2018.12.005

Fang, Y., Chen, J., Cho, Y. K., Kim, K., Zhang, S., & Perez, E. (2018). Vision-based load sway monitoring to improve crane safety in blind lifts. *Journal of Structural Integrity and Maintenance*, *3*(4), 233–242. doi: 10.1080/24705314.2018.1531348

Fang, Y., & Cho, Y. K. (2017). Effectiveness Analysis from a Cognitive Perspective for a Real-Time Safety Assistance System for Mobile Crane Lifting Operations. *Journal of Construction Engineering and Management*, *143*(4). doi: 10.1061/(asce)co.1943-7862.0001258

Fareh, R., Baziyad, M., Rabie, T., & Bettayeb, M. (2020). Enhancing path quality of real-time path planning algorithms for mobile robots: A sequential linear paths approach. *IEEE Access*, *8*, 167090–167104. doi: 10.1109/ACCESS.2020.3016525

Fatemi, A., & Vangt, L. (1998). Cumulative fatigue damage and life prediction theories: a survey of the state of the art for homogeneous materials. In Int. 1. Fatigue (Vol. 20, Issue I).

Ferguson, D., Kalra, N., & Stentz, A. (n.d.). *Replanning with RRTs*.

Ferguson, D., & Stentz, A. (n.d.). *Field D*: An Interpolation-Based Path Planner and Replanner*.

Fleet, D., Pajdla, T., Schiele, B., & Tuytelaars, T. (Eds.). (2014). *Computer Vision – ECCV 2014* (Vol. 8693). Cham: Springer International Publishing. doi: 10.1007/978-3-319-10602-1

Fung, I. W. H., Lo, T. Y., & Tung, K. C. F. (2012). Towards a better reliability of risk assessment: Development of a qualitative & quantitative risk evaluation model (Q 2REM) for different trades of construction works in Hong Kong. *Accident Analysis and Prevention*, *48*, 167–184. doi: 10.1016/j.aap.2011.05.011

Gao, J., Geng, X., Zhang, Y., Wang, R., & Shao, K. (2024). Augmented weighted bidirectional feature pyramid network for marine object detection. *Expert Systems with Applications*, *237*. doi: 10.1016/j.eswa.2023.121688

Garrett, C. R., Chitnis, R., Holladay, R., Kim, B., Silver, T., Kaelbling, L. P., & Lozano-Pérez, T. (2021). Annual Review of Control, Robotics, and Autonomous Systems Integrated

Task and Motion Planning. *Annu. Rev. Control Robot. Auton. Syst*, *4*, 265–293. doi: 10.1146/annurev-control-091420

Gashaw, T., & Jilcha, K. (2022). Developing a fuzzy synthetic evaluation model for risk assessment: a case of Addis-Djibouti railway construction project. *Innovative Infrastructure Solutions*, *7*(2). doi: 10.1007/s41062-022-00753-8

Goutham, M., Boyle, S., Menon, M., Mohan, S., Garrow, S., & Stockar, S. S. (2023). Optimal Path Planning Through a Sequence of Waypoints. *IEEE Robotics and Automation Letters*, *8*(3), 1509–1514. doi: 10.1109/LRA.2023.3240662

Gu, B., Guo, H., Huang, Y., Lim, H. W., & Fang, D. (2022). Computer vision-based human-machine collision warning system on construction site. *IOP Conference Series: Earth and Environmental Science*, *1101*(3). doi: 10.1088/1755-1315/1101/3/032014

Gu, J., Qin, Y., Xia, Y., Wang, J., Gao, H., & Jiao, Q. (2021). Failure Analysis and Prevention for Tower Crane as Sudden Unloading. *Journal of Failure Analysis and Prevention*, *21*(5), 1590–1595. doi: 10.1007/s11668-021-01201-y

Guerra-Fuentes, L., Torres-López, M., Hernandez-Rodriguez, M. A. L., & Garcia-Sanchez, E. (2020). Failure analysis of steel wire rope used in overhead crane system. *Engineering Failure Analysis*, *118*. doi: 10.1016/j.engfailanal.2020.104893

Gul, F., Mir, I., & Mir, S. (2023). Aquila Optimizer with parallel computing strategy for efficient environment exploration. *Journal of Ambient Intelligence and Humanized Computing*, *14*(4), 4175–4190. doi: 10.1007/s12652-023-04515-x

Gul, F., & Rahiman, W. (2019). An Integrated approach for Path Planning for Mobile Robot Using Bi-RRT. *IOP Conference Series: Materials Science and Engineering*, *697*(1). doi: 10.1088/1757-899X/697/1/012022

Gunduz, M., & Laitinen, H. (2018). Construction safety risk assessment with introduced control levels. *Journal of Civil Engineering and Management*, *24*(1), 11–18. doi: 10.3846/jcem.2018.284

Guo, H., Zhou, Y., Pan, Z., & Lin, X. (2021). Automated lift planning methods for mobile cranes. In Automation in Construction (Vol. 132). Elsevier B.V. doi: 10.1016/j.autcon.2021.103982

Guo, P., Ma, X., Zhang, W., Gao, F., Liang, M., & Shi, H. (2021). Study of Fault Detection of Bridge Crane Wheel based on Fourier Transform. *ACM International Conference Proceeding Series*, 23–27. doi: 10.1145/3458359.3458379

Gürcanli, G. E., Baradan, S., & Uzun, M. (2015). Risk perception of construction equipment operators on construction sites of Turkey. *International Journal of Industrial Ergonomics*, *46*, 59–68. doi: 10.1016/j.ergon.2014.12.004

Halim, N. A. B. A., & Ismail, A. W. B. (2021). ARHome: Object Selection and Manipulation using Raycasting Technique with 3D-model Tracking in Handheld Augmented Reality. *2021 IEEE 6th International Conference on Computing, Communication and Automation, ICCCA 2021*, 254–259. doi: 10.1109/ICCCA52192.2021.9666416

He, F., Ong, S. K., & Nee, A. Y. C. (2021). An integrated mobile augmented reality digital twin monitoring system. *Computers*, *10*(8). doi: 10.3390/computers10080099

Henrich, D. (1996). *A review of parallel processing approaches to motion planning*.

Henrich, D. (1997). Fast Motion Planning by Parallel Processing-A Review. In Journal of Intelligent and Robotic Systems (Vol. 20). Kluwer Academic Publishers.

Hinze, J., Huang, X., & Terry, L. (2005). The Nature of Struck-by Accidents. *Journal of Construction Engineering and Management*, *131*(2), 262–268. doi: 10.1061/(asce)0733-9364(2005)131:2(262)

Hinze, J., Member, /, Pedersen, C., & Fredley3, J. (n.d.). *IDENTIFYING ROOT CAUSES OF CONSTRUCTION INJURIES*.

Hou, X., Li, C., & Fang, Q. (2023). Computer vision-based safety risk computing and visualization on construction sites. *Automation in Construction*, *156*. doi: 10.1016/j.autcon.2023.105129

Hu, S., & Fang, Y. (n.d.). *Automating Crane Lift Path through Integration of BIM and Path Finding Algorithm* (Vol. 2020).

Hu, S., Fang, Y., & Bai, Y. (2021). Automation and optimization in crane lift planning: A critical review. *Advanced Engineering Informatics*, *49*. doi: 10.1016/j.aei.2021.101346

Hung, P. D., & Su, N. T. (2021a). Unsafe Construction Behavior Classification Using Deep Convolutional Neural Network. *Pattern Recognition and Image Analysis*, *31*(2), 271–284. doi: 10.1134/S1054661821020073

Hung, P. D., & Su, N. T. (2021b). Unsafe Construction Behavior Classification Using Deep Convolutional Neural Network. *Pattern Recognition and Image Analysis*, *31*(2), 271–284. doi: 10.1134/S1054661821020073

Hussain, M., Ye, Z., Chi, H. L., & Hsu, S. C. (2024). Predicting degraded lifting capacity of aging tower cranes: A digital twin-driven approach. *Advanced Engineering Informatics*, *59*. doi: 10.1016/j.aei.2023.102310

Hussain, M., Zheng, B., Chi, H. L., & Hsu, S. C. (2023). Automated and continuous BIM-based life cycle carbon assessment for infrastructure design projects. *Resources, Conservation and Recycling*, *190*. doi: 10.1016/j.resconrec.2022.106848

Ismail, A. W., Halim, N. A. A., Talib, R., & Sihes, A. J. (2023). Target selection method on the occluded and distant object in handheld augmented reality. *Indonesian Journal of Electrical Engineering and Computer Science*, *29*(2), 1157–1165. doi: 10.11591/ijeecs.v29.i2.pp1157-1165

ISSH. (2024). *Occupational injuries in construction in 2022 - ISSH07/2024 Statistical Highlights*. Retrieved from www.legco.gov.hk

Jannadi, O. A., & Almishari, S. (n.d.). *Risk Assessment in Construction*. doi: 10.1061/ASCE0733-93642003129:5492

Jeelani, I., Asce, S. M., Albert, A., Asce, A. M., Gambatese, J. A., & Asce, M. (2016). *Why Do Construction Hazards Remain Unrecognized at the Work Interface?* doi: 10.1061/(ASCE)

Ji, Y., & Leite, F. (2018). Automated tower crane planning: leveraging 4-dimensional BIM and rule-based checking. *Automation in Construction*, *93*, 78–90. doi: 10.1016/j.autcon.2018.05.003

Jia, W., Wang, W., & Zhang, Z. (2023). From simple digital twin to complex digital twin part II: Multi-scenario applications of digital twin shop floor. *Advanced Engineering Informatics*, *56*. doi: 10.1016/j.aei.2023.101915

Jiang, F., Ding, Y., Song, Y., Geng, F., & Wang, Z. (2021). Digital Twin-driven framework for fatigue life prediction of steel bridges using a probabilistic multiscale model: Application to segmental orthotropic steel deck specimen. *Engineering Structures*, *241*. doi: 10.1016/j.engstruct.2021.112461

Jiang, W., & Ding, L. (2024a). Unsafe hoisting behavior recognition for tower crane based on transfer learning. *Automation in Construction*, *160*. doi: 10.1016/j.autcon.2024.105299

Jiang, W., & Ding, L. (2024b). Unsafe hoisting behavior recognition for tower crane based on transfer learning. *Automation in Construction*, *160*. doi: 10.1016/j.autcon.2024.105299

Jiang, W., Ding, L., & Zhou, C. (2022). Digital twin: Stability analysis for tower crane hoisting safety with a scale model. *Automation in Construction*, *138*. doi: 10.1016/j.autcon.2022.104257

Jiang, Y., Li, M., Wu, W., Wu, X., Zhang, X., Huang, X., Zhong, R. Y., & Huang, G. G. Q. (2023). Multi-domain ubiquitous digital twin model for information management of complex infrastructure systems. *Advanced Engineering Informatics*, *56*. doi: 10.1016/j.aei.2023.101951

Jiao, X., Wu, N., Zhang, X., Fan, J., Cai, Z., Wang, Y., & Zhou, Z. (2024). Enhancing Tower Crane Safety: A UAV-Based Intelligent Inspection Approach. *Buildings*, *14*(5). doi: 10.3390/buildings14051420

Johnson, C. (2002). *Software tools to support incident reporting in safety-critical systems*. Retrieved from www.elsevier.com/locate/ssci

Jung, H., Choi, B., Kang, S., & Kang, Y. (2022). Temporal analysis of the frequency of accidents associated with construction equipment. *Safety Science*, *153*. doi: 10.1016/j.ssci.2022.105817

Kalairassan, G., Boopathi, M., & Mohan, R. M. (2017). Analysis of load monitoring system in hydraulic mobile cranes. *IOP Conference Series: Materials Science and Engineering*, *263*(6). doi: 10.1088/1757-899X/263/6/062045

Kasapakis, V., & Gavalas, D. (2017). Occlusion handling in outdoors augmented reality games. *Multimedia Tools and Applications*, *76*(7), 9829–9854. doi: 10.1007/s11042-016-3581-1

Kayhani, N., Taghaddos, H., Mousaei, A., Behzadipour, S., & Hermann, U. (2021). Heavy mobile crane lift path planning in congested modular industrial plants using a robotics approach. *Automation in Construction*, *122*. doi: 10.1016/j.autcon.2020.103508

Kazemian, A., Yuan, X., Davtalab, O., & Khoshnevis, B. (2019). Computer vision for real-time extrusion quality monitoring and control in robotic construction. *Automation in Construction*, *101*, 92–98. doi: 10.1016/j.autcon.2019.01.022

Khallaf, R., & Khallaf, M. (2021). Classification and analysis of deep learning applications in construction: A systematic literature review. *Automation in Construction*, *129*. doi: 10.1016/j.autcon.2021.103760

Khowaja, S. A., Khuwaja, P., Dharejo, F. A., Raza, S., Lee, I. H., Naqvi, R. A., & Dev, K. (2024). ReFuSeAct: Representation fusion using self-supervised learning for activity recognition in next generation networks. *Information Fusion*, *102*. doi: 10.1016/j.inffus.2023.102044

Kim, B., An, E. J., Kim, S., Sri Preethaa, K. R., Lee, D. E., & Lukacs, R. R. (2024). SRGAN-enhanced unsafe operation detection and classification of heavy construction machinery using cascade learning. *Artificial Intelligence Review*, *57*(8). doi: 10.1007/s10462-024-10839-7

Kim, H., Kim, H., Hong, Y. W., & Byun, H. (2018). Detecting Construction Equipment Using a Region-Based Fully Convolutional Network and Transfer Learning. *Journal of Computing in Civil Engineering*, *32*(2). doi: 10.1061/(asce)cp.1943-5487.0000731

Kiss, B., & Szalay, Z. (2020). Modular approach to multi-objective environmental optimization of buildings. *Automation in Construction*, *111*. doi: 10.1016/j.autcon.2019.103044

Koenig, S., Likhachev, M., & Furcy, D. (2004). Lifelong Planning A*. *Artificial Intelligence*, *155*(1–2), 93–146. doi: 10.1016/j.artint.2003.12.001

Kolar, Z., Chen, H., & Luo, X. (2018). Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images. *Automation in Construction*, *89*, 58–70. doi: 10.1016/j.autcon.2018.01.003

Kong, F., Lu, Z., Kong, L., & Chen, T. (2023). Information field in a manufacturing System: Concepts, measurements and applications. In Advanced Engineering Informatics (Vol. 56). Elsevier Ltd. doi: 10.1016/j.aei.2023.101946

Kulinan, A. S., Park, M., Aung, P. P. W., Cha, G., & Park, S. (2024). Advancing construction site workforce safety monitoring through BIM and computer vision integration. *Automation in Construction*, *158*. doi: 10.1016/j.autcon.2023.105227

Kulka, J., Mantic, M., Faltinova, E., Molnar, V., & Fedorko, G. (2018). Failure analysis of the foundry crane to increase its working parameters. *Engineering Failure Analysis*, *88*, 25–34. doi: 10.1016/j.engfailanal.2018.02.020

Kulka, J., Mantic, M., Fedorko, G., & Molnar, V. (2016). Analysis of crane track degradation due to operation. *Engineering Failure Analysis*, *59*, 384–395. doi: 10.1016/j.engfailanal.2015.11.009

Kušić, K., Schumann, R., & Ivanjko, E. (2023). A digital twin in transportation: Real-time synergy of traffic data streams and simulation for virtualizing motorway dynamics. *Advanced Engineering Informatics*, *55*. doi: 10.1016/j.aei.2022.101858

Lan, Q., Zhang, D., & Li, Y. (2017). *Analysis on the Current Situation and Countermeasures of Elevator Safety in China*.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi Twitter, W. (2016). *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*.

Lee, H.-S., Kim, H., Park, M., Ai Lin Teo, E., & Lee, K.-P. (2012). Construction Risk Assessment Using Site Influence Factors. *Journal of Computing in Civil Engineering*, *26*(3), 319–330. doi: 10.1061/(asce)cp.1943-5487.0000146

Lee, J. G., Hwang, J., Chi, S., & Seo, J. (2022). Synthetic Image Dataset Development for Vision-Based Construction Equipment Detection. *Journal of Computing in Civil Engineering*, *36*(5). doi: 10.1061/(asce)cp.1943-5487.0001035

Lee, J., Phillips, I., & Lynch, Z. (2020). Causes and prevention of mobile crane-related accidents in South Korea. *International Journal of Occupational Safety and Ergonomics*, 1–10. doi: 10.1080/10803548.2020.1775384

Lee, W. J., Wu, H., Yun, H., Kim, H., Jun, M. B. G., & Sutherland, J. W. (2019). Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data. *Procedia CIRP*, *80*, 506–511. doi: 10.1016/j.procir.2018.12.019

Lei, Z., Taghaddos, H., Olearczyk, J., Al-Hussein, M., & Hermann, U. (2013). Automated Method for Checking Crane Paths for Heavy Lifts in Industrial Projects. *Journal of Construction Engineering and Management*, *139*(10). doi: 10.1061/(asce)co.1943-7862.0000740

Li, A. (2024). Human error risk prioritization in crane operations based on CPT and ICWGT. *PLoS ONE*, *19*(2 February). doi: 10.1371/journal.pone.0297120

Li, W., Li, H., Wang, Y., & Han, Y. (2024). Optimizing flexible job shop scheduling with automated guided vehicles using a multi-strategy-driven genetic algorithm. *Egyptian Informatics Journal*, *25*. doi: 10.1016/j.eij.2023.100437

Li, X., Chi, H. lin, Wu, P., & Shen, G. Q. (2020). Smart work packaging-enabled constraint-free path re-planning for tower crane in prefabricated products assembly process. *Advanced Engineering Informatics*, *43*. doi: 10.1016/j.aei.2019.101008

Li, X., Chi, H.-L., & Shen, G. Q. (n.d.). *Constraint-free Crane Path Re-Planning*.

Li, Yibo, Dong, D., & Guo, X. (2020). *Mobile robot path planning based on improved genetic algorithm with A-star heuristic method*. 1306–1311. doi: 10.1109/ITAIC49862.2020.9338968

Li, Yue, Zhao, J., Chen, Z., Xiong, G., & Liu, S. (2023). A Robot Path Planning Method Based on Improved Genetic Algorithm and Improved Dynamic Window Approach. *Sustainability (Switzerland)*, *15*(5). doi: 10.3390/su15054656

Liang, Y., & Liu, Q. (2022). Early warning and real-time control of construction safety risk of underground engineering based on building information modeling and internet of things. *Neural Computing and Applications*, *34*(5), 3433–3442. doi: 10.1007/s00521-021-05755-8

Lin, K. L., & Fang, J. L. (2013). Applications of computer vision on tile alignment inspection. *Automation in Construction*, *35*, 562–567. doi: 10.1016/j.autcon.2013.01.009

Lin, Y., Wu, D., Wang, X., Wang, X., & Gao, S. (2014). Lift path planning for a nonholonomic crawler crane. *Automation in Construction*, *44*, 12–24. doi: 10.1016/j.autcon.2014.03.007

Linda Levine. (2008). *Worker Safety in the Construction Industry: The Crane and Derrick Standard*.

Liu, Jiajing, Fang, W., Love, P. E. D., Hartmann, T., Luo, H., & Wang, L. (2022). Detection and location of unsafe behaviour in digital images: A visual grounding approach. *Advanced Engineering Informatics*, *53*. doi: 10.1016/j.aei.2022.101688

Liu, Jiajing, Luo, H., & Liu, H. (2022). Deep learning-based data analytics for safety in construction. In Automation in Construction (Vol. 140). Elsevier B.V. doi: 10.1016/j.autcon.2022.104302

Liu, Jinfeng, Cao, X., Zhou, H., Li, L., Liu, X., Zhao, P., & Dong, J. (2021a). A digital twin-driven approach towards traceability and dynamic control for processing quality. *Advanced Engineering Informatics*, *50*. doi: 10.1016/j.aei.2021.101395

Liu, Jinfeng, Cao, X., Zhou, H., Li, L., Liu, X., Zhao, P., & Dong, J. (2021b). A digital twin-driven approach towards traceability and dynamic control for processing quality. *Advanced Engineering Informatics*, *50*. doi: 10.1016/j.aei.2021.101395

Liu, Jinfeng, Du, X., Zhou, H., Liu, X., Li, L. ei, & Feng, F. (2019). A digital twin-based approach for dynamic clamping and positioning of the flexible tooling system. *Procedia CIRP*, *80*, 746–749. doi: 10.1016/j.procir.2019.01.063

Liu, Jiping, Kang, X., Dong, C., & Zhang, F. (2019). *Simulation of Real-Time Path Planning for Large-Scale Transportation Network Using Parallel Computation*.

Liu, P., Chi, H. L., Li, X., & Guo, J. (2021). Effects of dataset characteristics on the performance of fatigue detection for crane operators using hybrid deep neural networks. *Automation in Construction*, *132*. doi: 10.1016/j.autcon.2021.103901

Liu, Q., Wang, C., Li, X., & Gao, L. (2023). An improved genetic algorithm with modified critical path-based searching for integrated process planning and scheduling problem considering automated guided vehicle transportation task. *Journal of Manufacturing Systems*, *70*, 127–136. doi: 10.1016/j.jmsy.2023.07.004

Liu, W., Meng, Q., Li, Z., & Hu, X. (2021). Applications of computer vision in monitoring the unsafe behavior of construction workers: Current status and challenges. In Buildings (Vol. 11, Issue 9). MDPI. doi: 10.3390/buildings11090409

Liu, X., Jiang, D., Tao, B., Xiang, F., Jiang, G., Sun, Y., Kong, J., & Li, G. (2023). A systematic review of digital twin about physical entities, virtual models, twin data, and applications. *Advanced Engineering Informatics*, *55*, 101876. doi: 10.1016/j.aei.2023.101876

Luan, P. G., & Thinh, N. T. (2023). Hybrid genetic algorithm based smooth global-path planning for a mobile robot. *Mechanics Based Design of Structures and Machines*, *51*(3), 1758–1774. doi: 10.1080/15397734.2021.1876569

Lynch, K. (Kevin M.), & Park, F. C. (n.d.). *Modern robotics : mechanics, planning, and control*.

Lyu, H.-M., Sun, W.-J., Shen, S.-L., & Zhou, A.-N. (2020). Risk Assessment Using a New Consulting Process in Fuzzy AHP. *Journal of Construction Engineering and Management*, *146*(3). doi: 10.1061/(asce)co.1943-7862.0001757

Madhu, K. A., Rajakumara, H. N., Kangda, M. Z., Girija, S. P., Reddy, G. R., & Wodajo, A. W. (2023). Evaluation of optimal route selection for public transport network routes on urban roads using Fuzzy-TOPSIS method . *Engineering Reports*. doi: 10.1002/eng2.12794

Martin, A., Hill, A. J., Seiler, K. M., & Balamurali, M. (2023). Automatic excavator action recognition and localisation for untrimmed video using hybrid LSTM-Transformer networks. *International Journal of Mining, Reclamation and Environment*. doi: 10.1080/17480930.2023.2290364

Mdot. (n.d.). *DEVELOPMENT OF STEEL BEAM END DETERIORATION GUIDELINES FINAL REPORT-JANUARY 2005 CENTER FOR STRUCTURAL DURABILITY MICHIGAN TECH TRANSPORTATION INSTITUTE*.

Miao, H., & Tian, Y. C. (2013). Dynamic robot path planning using an enhanced simulated annealing approach. *Applied Mathematics and Computation*, *222*, 420–437. doi: 10.1016/j.amc.2013.07.022

Mohajer, B., Kiani, K., Samiei, E., & Sharifi, M. (2013). A new online random particles optimization algorithm for mobile robot path planning in dynamic environments. *Mathematical Problems in Engineering*, *2013*. doi: 10.1155/2013/491346

Moragane, H. P. M. N. L. B., Perera, B. A. K. S., Palihakkara, A. D., & Ekanayake, B. (2024). Application of computer vision for construction progress monitoring: a qualitative investigation. *Construction Innovation*, *24*(2), 446–469. doi: 10.1108/CI-05-2022-0130

Mustafa, F. E., Ahmed, I., Basit, A., Alvi, U. E. H., Malik, S. H., Mahmood, A., & Ali, P. R. (2023). A review on effective alarm management systems for industrial process control: Barriers and opportunities. In International Journal of Critical Infrastructure Protection (Vol. 41). Elsevier B.V. doi: 10.1016/j.ijcip.2023.100599

Nath, N. D., Behzadan, A. H., & Paal, S. G. (2020). Deep learning for site safety: Real-time detection of personal protective equipment. *Automation in Construction*, *112*. doi: 10.1016/j.autcon.2020.103085

Nearchou, A. C., & Aspragathos, N. A. (1997). A genetic path planning algorithm for redundant articulated robots. *Robotica*, *15*(2), 213–224. doi: 10.1017/S0263574797000234

Neitzel, R. L., Seixas, N. S., & Ren, K. K. (2001). A Review of Crane Safety in the Construction Industry. *Applied Occupational and Environmental Hygiene*, *16*(12), 1106–1117. doi: 10.1080/10473220127411

Niu, Y., Lu, W., Chen, K., Huang, G. G., & Anumba, C. (2016). Smart Construction Objects. *Journal of Computing in Civil Engineering*, *30*(4). doi: 10.1061/(asce)cp.1943-5487.0000550

Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., Assran, M., Ballas, N., Galuba, W., Howes, R., Huang, P.-Y., Li, S.-W., Misra, I., Rabbat, M., Sharma, V., … Bojanowski, P. (2023). *DINOv2: Learning Robust Visual Features without Supervision*. Retrieved from http://arxiv.org/abs/2304.07193

Pal, U., Mukhopadhyay, G., Sharma, A., & Bhattacharya, S. (2018). Failure analysis of wire rope of ladle crane in steel making shop. *International Journal of Fatigue*, *116*, 149–155. doi: 10.1016/j.ijfatigue.2018.06.019

Pan, W., & Hon, C. K. (2020). Briefing: Modular integrated construction for high-rise buildings. *Proceedings of the Institution of Civil Engineers: Municipal Engineer*, *173*(2), 64–68. doi: 10.1680/jmuen.18.00028

Panda, M., Das, B., Subudhi, B., & Pati, B. B. (2020). A Comprehensive Review of Path Planning Algorithms for Autonomous Underwater Vehicles. In International Journal of Automation and Computing (Vol. 17, Issue 3, pp. 321–352). Chinese Academy of Sciences. doi: 10.1007/s11633-019-1204-9

Pandey, M., Mishra, R., & Khare, A. (2024). Vision-based Human Activity Recognition Using Local Phase Quantization. *Journal of Artificial Intelligence and Technology*. doi: 10.37965/jait.2024.0351

Paneru, S., & Jeelani, I. (2021). Computer vision applications in construction: Current state, opportunities & challenges. In Automation in Construction (Vol. 132). Elsevier B.V. doi: 10.1016/j.autcon.2021.103940

Pantic, M., Meijer, I., Bahnemann, R., Alatur, N., Andersson, O., Cadena, C., Siegwart, R., & Ott, L. (2023). Obstacle avoidance using Raycasting and Riemannian Motion Policies at kHz rates for MAVs. *Proceedings - IEEE International Conference on Robotics and Automation*, *2023-May*, 1666–1672. doi: 10.1109/ICRA48891.2023.10161365

Pape, J. A., & Neu, R. W. (2001). Fretting fatigue damage accumulation in PH13-8Mo stainless steel. In International Journal of Fatigue (Vol. 23). Retrieved from www.elsevier.com/locate/ijfatigue

Paulin, G., Sambolek, S., & Ivasic-Kos, M. (2021, September 8). Person localization and distance determination using the raycast method. *2021 6th International Conference on Smart and Sustainable Technologies, SpliTech 2021*. doi: 10.23919/SpliTech52315.2021.9566329

Pauwels, P., de Koning, R., Hendrikx, B., & Torta, E. (2023). Live semantic data from building digital twins for robot navigation: Overview of data transfer methods. *Advanced Engineering Informatics*, *56*. doi: 10.1016/j.aei.2023.101959

Pompetzki, M. A., Topper, T. H., & Duquesnay, D. L. (1990). The effect of compressive underloads and tensile overloads on fatigue damage accumulation in SAE 1045 steel. In Int J Fatigue (Vol. 12).

Precup, R. E., Hedrea, E. L., Roman, R. C., Petriu, E. M., Bojan-Dragos, C. A., Szedlak-Stinean, A. I., & Paulescu, F. C. (2022). AVOA-Based Tuning of Low-Cost Fuzzy Controllers for Tower Crane Systems. *IEEE International Conference on Fuzzy Systems*, *2022-July*. doi: 10.1109/FUZZ-IEEE55066.2022.9882748

Priyanka, E. B., & Thangavel, S. (2022). Multi-type feature extraction and classification of leakage in oil pipeline network using digital twin technology. *Journal of Ambient Intelligence and Humanized Computing*. doi: 10.1007/s12652-022-03818-9

Przybyłek, G., & Więckowski, J. (2022). Method of assessing the technical condition and failure of overhead cranes designed to work in difficult conditions. *Case Studies in Construction Materials*, *16*. doi: 10.1016/j.cscm.2021.e00811

Rahim Abdul Hamid, A., Azhari, R., Zakaria, R., Aminudin, E., Putra Jaya, R., Nagarajan, L., Yahya, K., Haron, Z., & Yunus, R. (2019). Causes of crane accidents at construction sites in Malaysia. *IOP Conference Series: Earth and Environmental Science*, *220*(1). doi: 10.1088/1755-1315/220/1/012028

Raj, A., & Teizer, J. (2024, May 27). State of the Art Review of Technological Advancements for Safe Tower Crane Operation. *Proceedings of the 41st International Symposium on Automation and Robotics in Construction*. doi: 10.22260/isarc2024/0048

Rayco Wylie. (2012). *R147 WIRELESS ANTI-TWO-BLOCK INDICATOR Installation and Operation Manual*. Retrieved from www.craneindicators.com

Razali, M. R., Mohd Faudzi, A. A., Shamsudin, A. U., & Mohamaddan, S. (2023). A hybrid controller method with genetic algorithm optimization to measure position and angular for mobile robot motion control. *Frontiers in Robotics and AI*, *9*. doi: 10.3389/frobt.2022.1087371

Reja, V. K., Varghese, K., & Ha, Q. P. (2022). Computer vision-based construction progress monitoring. In Automation in Construction (Vol. 138). Elsevier B.V. doi: 10.1016/j.autcon.2022.104245

Roberge, V., Tarbouchi, M., & Labonte, G. (2013). Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning. *IEEE Transactions on Industrial Informatics*, *9*(1), 132–141. doi: 10.1109/TII.2012.2198665

Roman, R. C., Precup, R. E., & Petriu, E. M. (2021). Hybrid data-driven fuzzy active disturbance rejection control for tower crane systems. *European Journal of Control*, *58*, 373–387. doi: 10.1016/j.ejcon.2020.08.001

Roman, R. C., Precup, R. E., Petriu, E. M., & Dragan, F. (2019). Combination of data-driven active disturbance rejection and Takagi-Sugeno fuzzy control with experimental validation on tower crane systems. *Energies*, *12*(8). doi: 10.3390/en12081548

Roysson, S., Gustafsson, J., Lindell, R., & Sitompul, T. A. (n.d.). *EVALUATING THE LIFTING CAPACITY IN A MOBILE CRANE SIMULATION*.

Sabuhi, M., Zhou, M., Bezemer, C. P., & Musilek, P. (2021). Applications of Generative Adversarial Networks in Anomaly Detection: A Systematic Literature Review. In IEEE Access (Vol. 9, pp. 161003–161029). Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ACCESS.2021.3131949

Sadeghi, H., & Zhang, X. (2024). Towards safer tower crane operations: An innovative knowledge-based decision support system for automated safety risk assessment. *Journal of Safety Research*, *90*, 272–294. doi: 10.1016/j.jsr.2024.05.011

Sadeghi, H., Zhang, X., & Mohandes, S. R. (2023). Developing an ensemble risk analysis framework for improving the safety of tower crane operations under coupled Fuzzy-based environment. *Safety Science*, *158*. doi: 10.1016/j.ssci.2022.105957

Sadeghi, S., Soltanmohammadlou, N., & Rahnamayiezekavat, P. (2021). A systematic review of scholarly works addressing crane safety requirements. *Safety Science*, *133*. doi: 10.1016/j.ssci.2020.105002

Sanni-Anibire, M. O., Mahmoud, A. S., Hassanain, M. A., & Salami, B. A. (2020). A risk assessment approach for enhancing construction safety performance. *Safety Science*, *121*, 15–29. doi: 10.1016/j.ssci.2019.08.044

Sarraf, R., & McGuire, M. P. (2020). Integration and comparison of multi-criteria decision making methods in safe route planner. *Expert Systems with Applications*, *154*. doi: 10.1016/j.eswa.2020.113399

Schieg, M. (2006). Risk management in construction project management. *Journal of Business Economics and Management*, *7*(2), 77–83. doi: 10.1080/16111699.2006.9636126

Seo, J., Han, S., Lee, S., & Kim, H. (2015). Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics*, *29*(2), 239–251. doi: 10.1016/j.aei.2015.02.001

Sepasgozar, S. M. E. (2021). Differentiating digital twin from digital shadow: Elucidating a paradigm shift to expedite a smart, sustainable built environment. In Buildings (Vol. 11, Issue 4). MDPI AG. doi: 10.3390/buildings11040151

Shafique, M., & Rafiq, M. (2019). An overview of construction occupational accidents in Hong Kong: A recent trend and future perspectives. *Applied Sciences (Switzerland)*, *9*(10). doi: 10.3390/app9102069

Shaikh, A. A., & Kumar, D. D. (2016). Lifting capacity enhancement of a crawler crane by improving stability. *Journal of Theoretical and Applied Mechanics (Poland)*, *54*(1), 219–227. doi: 10.15632/jtam-pl.54.1.219

Shao, B., Hu, Z., Liu, Q., Chen, S., & He, W. (2019). Fatal accident patterns of building construction activities in China. *Safety Science*, *111*, 253–263. doi: 10.1016/j.ssci.2018.07.019

Shapira, A., Asce, F., & Lyachin, B. (n.d.). *Identification and Analysis of Factors Affecting Safety on Construction Sites with Tower Cranes*. doi: 10.1061/ASCE0733-93642009135:124

Shen, Xingwang, Li, X., Zhou, B., Jiang, Y., & Bao, J. (2023). Dynamic knowledge modeling and fusion method for custom apparel production process based on knowledge graph. *Advanced Engineering Informatics*, *55*. doi: 10.1016/j.aei.2023.101880

Shen, Xu, Asce, S. M., Marks, E., Pradhananga, N., & Cheng, T. (2016). *Hazardous Proximity Zone Design for Heavy Construction Excavation Equipment*. doi: 10.1061/(ASCE)CO.1943-7862

Shi, L., Mehrooz, G., & Jacobsen, R. H. (2021). Inspection Path Planning for Aerial Vehicles via Sampling-based Sequential Optimization. *2021 International Conference on Unmanned Aircraft Systems, ICUAS 2021*, 679–687. doi: 10.1109/ICUAS51884.2021.9476784

Shin, I. J. (2015a). Factors that affect safety of tower crane installation/dismantling in construction industry. *Safety Science*, *72*, 379–390. doi: 10.1016/j.ssci.2014.10.010

Shin, I. J. (2015b). Factors that affect safety of tower crane installation/dismantling in construction industry. *Safety Science*, *72*, 379–390. doi: 10.1016/j.ssci.2014.10.010

Shin, Y., Choi, Y., Won, J., Hong, T., & Koo, C. (2024). A New Benchmark Model for the Automated Detection and Classification of a Wide Range of Heavy Construction Equipment. *Journal of Management in Engineering*, *40*(2). doi: 10.1061/jmenea.meeng-5630

Singh, A., Rana, A. J., Kumar, A., Vyas, S., & Singh Rawat, Y. (2024). *Semi-supervised Active Learning for Video Action Detection*. Retrieved from www.aaai.org

Six, K., Mihalj, T., Trummer, G., Marte, C., Krishna, V. v., Hossein-Nia, S., & Stichel, S. (2020). Assessment of running gear performance in relation to rolling contact fatigue of wheels and rails based on stochastic simulations. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *234*(4), 405–416. doi: 10.1177/0954409719879600

Soltani, A. R., Tawfik, H., Goulermas, J. Y., & Fernando, T. (2002a). Path planning in construction sites: Performance evaluation of the dijkstra, a*, and GA search algorithms. *Advanced Engineering Informatics*, *16*(4), 291–303. doi: 10.1016/S1474-0346(03)00018-1

Soltani, A. R., Tawfik, H., Goulermas, J. Y., & Fernando, T. (2002b). Path planning in construction sites: Performance evaluation of the dijkstra, a*, and GA search algorithms. *Advanced Engineering Informatics*, *16*(4), 291–303. doi: 10.1016/S1474-0346(03)00018-1

Soltani, A. R., Tawfik, H., Goulermas, J. Y., & Fernando, T. (2002c). Path planning in construction sites: Performance evaluation of the dijkstra, a*, and GA search algorithms. *Advanced Engineering Informatics*, *16*(4), 291–303. doi: 10.1016/S1474-0346(03)00018-1

Son, H., & Kim, C. (2021). Integrated worker detection and tracking for the safe operation of construction machinery. *Automation in Construction*, *126*. doi: 10.1016/j.autcon.2021.103670

Srivastava, A. K., Kontoudis, G. P., Sofge, D., Otte, M., & Srivastava, A. K. (2022). *Distributed Multi-Robot Information Gathering using Path-Based Sensors in Entropy-Weighted Voronoi Regions*. Retrieved from https://www.researchgate.net/publication/363856270

Stein, M. (2018). Conducting Safety Inspections of Container Gantry Cranes Using Unmanned Aerial Vehicles. In Lecture Notes in Logistics (pp. 154–161). Springer Science and Business Media B.V. doi: 10.1007/978-3-319-74225-0_20

Stentz, A. (1994). *Optimal and Efficient Path Planning for Partially-Known Environments*.

Stentz, A. (1995). *The Focussed D* Algorithm for Real-Time Replanning*.

Steven Chun-yin, H. H., Hon KWOK Wai-keung, J., Hon SHIU Ka-fai, J., Hon CHU Kwok-keung Hon CHAU Siu-chung Hon YIU Pak-leung, J., Hon Dennis LEUNG Tsz-wing, M., Hon Sunny TAN Hon Lillian KWOK Ling-lai Hon Stephen WONG Yuen-shan Hon Kingsley WONG Kwok, M., & Hon NGAN Man-yu, J. (2022). *Legislative Council*.

Sulankivi, K., Kiviniemi, M., & Mäkelä, T. (2014). *BIM-based Site Layout and Safety Planning*. Retrieved from https://www.researchgate.net/publication/38289116

Taghaddos, H., Hermann, U., & Abbasi, A. B. (2018). Automated Crane Planning and Optimization for modular construction. *Automation in Construction*, *95*, 219–232. doi: 10.1016/j.autcon.2018.07.009

Taheri, S., Vincent, L., & Le-Roux, J. C. (2013). A new model for fatigue damage accumulation of austenitic stainless steel under variable amplitude loading. *Procedia Engineering*, *66*, 575–586. doi: 10.1016/j.proeng.2013.12.109

Tajeen, H., & Zhu, Z. (2014). Image dataset development for measuring construction equipment recognition performance. *Automation in Construction*, *48*, 1–10. doi: 10.1016/j.autcon.2014.07.006

Tam, V. W. Y., & Fung, I. W. H. (2011). Tower crane safety in the construction industry: A Hong Kong study. *Safety Science*, *49*(2), 208–215. doi: 10.1016/j.ssci.2010.08.001

Tansar, H., Duan, H. F., & Mark, O. (2023). A multi-objective decision-making framework for implementing green-grey infrastructures to enhance urban drainage system resilience. *Journal of Hydrology*, *620*. doi: 10.1016/j.jhydrol.2023.129381

Teizer, J., & Cheng, T. (2015). Proximity hazard indicator for workers-on-foot near miss interactions with construction equipment and geo-referenced hazard areas. *Automation in Construction*, *60*, 58–73. doi: 10.1016/j.autcon.2015.09.003

Teng, Z., Wu, H., Boller, C., & Starke, P. (2020). Thermodynamic entropy as a marker of high-cycle fatigue damage accumulation: Example for normalized SAE 1045 steel. *Fatigue and Fracture of Engineering Materials and Structures*, *43*(12), 2854–2866. doi: 10.1111/ffe.13303

Tian, Y., Wang, X., Yao, H., Chen, J., Wang, Z., & Yi, L. (2018). Occlusion handling using moving volume and ray casting techniques for augmented reality systems. *Multimedia Tools and Applications*, *77*(13), 16561–16578. doi: 10.1007/s11042-017-5228-2

Touzani, H., Hadj-Abdelkader, H., Seguy, N., Bouchafa, S., & Séguy, N. (2021). Multi-Robot Task Sequencing & Automatic Path Planning for Cycle Time Optimization: Application for Car Production Line. *IEEE Robotics and Automation Letters*, *6*(2), 1. doi: 10.1109/LRA.2021.3057011ï

Tran, V. T., Thom Pham, H., Yang, B. S., & Tien Nguyen, T. (2012). Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine. *Mechanical Systems and Signal Processing*, *32*, 320–330. doi: 10.1016/j.ymssp.2012.02.015

Tsai, C. C., Huang, H. C., & Chan, C. K. (2011). Parallel elite genetic algorithm and its application to global path planning for autonomous robot navigation. *IEEE Transactions on Industrial Electronics*, *58*(10), 4813–4821. doi: 10.1109/TIE.2011.2109332

Tubaileh, A. (2016). Working time optimal planning of construction site served by a single tower crane. *Journal of Mechanical Science and Technology*, *30*(6), 2793–2804. doi: 10.1007/s12206-016-0346-8

Tuncer, A., & Yildirim, M. (2012). Dynamic path planning of mobile robots with improved genetic algorithm. *Computers and Electrical Engineering*, *38*(6), 1564–1572. doi: 10.1016/j.compeleceng.2012.06.016

Tylman, W., Kolczyński, J., & Anders, G. J. (2010). Fully automatic AI-based leak detection system. *Energy*, *35*(9), 3838–3848. doi: 10.1016/j.energy.2010.05.038

Unit II ISO/OSI Model in Communication Networks. (n.d.).

Vo, A. V., Laefer, D. F., & Byrne, J. (2021). Optimizing urban LiDAR flight path planning using a genetic algorithm and a dual parallel computing framework. *Remote Sensing*, *13*(21). doi: 10.3390/rs13214437

Vu, M. N., Lobe, A., Beck, F., Weingartshofer, T., Hartl-Nesic, C., & Kugi, A. (2022). Fast trajectory planning and control of a lab-scale 3D gantry crane for a moving target in an environment with obstacles. *Control Engineering Practice*, *126*. doi: 10.1016/j.conengprac.2022.105255

Vukelic, G., Pastorcic, D., & Vizentin, G. (2019). Failure analysis of a crane gear shaft. *Procedia Structural Integrity*, *18*, 406–412. doi: 10.1016/j.prostr.2019.08.182

Walbridge, S., Nik-Bakht, M., Tsun, K., Ng, W., Shome, M., Shahria, · M, Ashraf El Damatty, A. ·, & Lovegrove, G. (2022). *Lecture Notes in Civil Engineering*. Retrieved from https://link.springer.com/bookseries/15087

Wan, W., Harada, K., & Nagata, K. (2016). *Assembly Sequence Planning for Motion Planning*. Retrieved from http://arxiv.org/abs/1609.03108

Wang, J., Zhu, G., Wu, S., & Luo, C. (2021). Worker's Helmet Recognition and Identity Recognition Based on Deep Learning. *Open Journal of Modelling and Simulation*, *09*(02), 135–145. doi: 10.4236/ojmsi.2021.92009

Wei, J., Zhu, K., & Ye, Y. (2023). FanRay: Using Fan Mechanism in Ray-casting to Improve Target Selection in VR. *ACM International Conference Proceeding Series*, 1–6. doi: 10.1145/3604383.3604395

Weingartshofer, T., Bischof, B., Meiringer, M., Hartl-Nesic, C., & Kugi, A. (2023). Optimization-based path planning framework for industrial manufacturing

processes with complex continuous paths. *Robotics and Computer-Integrated Manufacturing*, *82*. doi: 10.1016/j.rcim.2022.102516

Wiethorn, J. D., Matthew, P. E., Gardiner, R., Anthony, P. E., Bond, E., Cox, P. E. E. P., & King, R. A. (n.d.). *TOWER CRANE LIFE EXPECTANCY AN EXAMINATION OF RECENT TRENDS TO ESTABLISH AGE LIMITS*.

Wong, J. M. W., Chiang, Y. H., & Ng, T. S. (2008). Construction and economic development: The case of Hong Kong. *Construction Management and Economics*, *26*(8), 815–826. doi: 10.1080/01446190802189927

Wong, J., Zhang, J., & Lee, J. (2015). *A Vision of the Future Construction Industry of Hong Kong*.

Woo, J. W., An, J. Y., Cho, M. G., & Kim, C. J. (2021). Integration of path planning, trajectory generation and trajectory tracking control for aircraft mission autonomy. *Aerospace Science and Technology*, *118*. doi: 10.1016/j.ast.2021.107014

Wu, B. J., Jin, L. H., Zheng, X. Z., & Chen, S. (2024). Coupling analysis of crane accident risks based on Bayesian network and the N-K model. *Scientific Reports*, *14*(1). doi: 10.1038/s41598-024-51425-9

Wu, B., Tang, Y., Li, Z., & Tang, K. (2021). Fatigue damage accumulation modelling of critical components subjected to moving crane loads in reinforced-concrete industrial buildings. *Engineering Failure Analysis*, *119*. doi: 10.1016/j.engfailanal.2020.104951

Wu, L., Leng, J., & Ju, B. (2021). Digital twins-based smart design and control of ultra-precision machining: A review. In Symmetry (Vol. 13, Issue 9). MDPI. doi: 10.3390/sym13091717

Wu, S., Hou, L., Zhang, G. (Kevin), & Chen, H. (2022). Real-time mixed reality-based visual warning for construction workforce safety. *Automation in Construction*, *139*. doi: 10.1016/j.autcon.2022.104252

Wu, Y., Zhou, L., Zheng, P., Sun, Y., & Zhang, K. (2022). A digital twin-based multidisciplinary collaborative design approach for complex engineering product development. *Advanced Engineering Informatics*, *52*. doi: 10.1016/j.aei.2022.101635

Xiao, B., & Kang, S.-C. (2020). *Development of an Image Data Set of Construction Machines for Deep Learning Object Detection*. doi: 10.1061/(ASCE)CP.1943

Xiao, B., & Kang, S.-C. (2021). Vision-Based Method Integrating Deep Learning Detection for Tracking Multiple Construction Machines. *Journal of Computing in Civil Engineering*, *35*(2). doi: 10.1061/(asce)cp.1943-5487.0000957

Xiao, B., Lin, Q., & Chen, Y. (2021). A vision-based method for automatic tracking of construction machines at nighttime based on deep learning illumination enhancement. *Automation in Construction*, *127*. doi: 10.1016/j.autcon.2021.103721

Xu, S., Wang, J., Shou, W., Ngo, T., Sadick, A. M., & Wang, X. (2021). Computer Vision Techniques in Construction: A Critical Review. *Archives of Computational Methods in Engineering*, *28*(5), 3383–3397. doi: 10.1007/s11831-020-09504-3

Xuehui, A., Li, Z., Zuguang, L., Chengzhi, W., Pengfei, L., & Zhiwei, L. (2021). Dataset and benchmark for detecting moving objects in construction sites. *Automation in Construction*, *122*. doi: 10.1016/j.autcon.2020.103482

Yahia, H. S., & Mohammed, A. S. (2023). Path planning optimization in unmanned aerial vehicles using meta-heuristic algorithms: a systematic review. *Environmental Monitoring and Assessment*, *195*(1). doi: 10.1007/s10661-022-10590-y

Yang Wang, I. P. W. S. and D. J. M. M. I. (2007). Mobile_Robot_Path_Planning_in_Dynamic_Environments. *2007 IEEE International Conference on Robotics and Automation*.

Yang, X., Ran, Y., Zhang, G., Wang, H., Mu, Z., & Zhi, S. (2022). A digital twin-driven hybrid approach for the prediction of performance degradation in transmission unit of CNC machine tool. *Robotics and Computer-Integrated Manufacturing*, *73*. doi: 10.1016/j.rcim.2021.102230

Yavari, H., Qajar, J., Aadnoy, B. S., & Khosravanian, R. (2023). Selection of Optimal Well Trajectory Using Multi-Objective Genetic Algorithm and TOPSIS Method. *Arabian Journal for Science and Engineering*. doi: 10.1007/s13369-023-08149-1

Zhang, C., & Hammad, A. (2012a). Improving lifting motion planning and re-planning of cranes with consideration for safety and efficiency. *Advanced Engineering Informatics*, *26*(2), 396–410. doi: 10.1016/j.aei.2012.01.003

Zhang, C., & Hammad, A. (2012b). Improving lifting motion planning and re-planning of cranes with consideration for safety and efficiency. *Advanced Engineering Informatics*, *26*(2), 396–410. doi: 10.1016/j.aei.2012.01.003

Zhang, L., Zhang, Y., & Li, Y. (2020). Path planning for indoor Mobile robot based on deep learning. *Optik*, *219*. doi: 10.1016/j.ijleo.2020.165096

Zhang, M., & Ge, S. (2022). *Vision and Trajectory-Based Dynamic Collision Prewarning Mechanism for Tower Cranes*. doi: 10.1061/(ASCE)

Zhang, W. J., Wan, H. P., Hu, P. H., Ge, H. Bin, Luo, Y., & Todd, M. D. (2024). Semi-supervised learning approach for construction object detection by integrating super-

resolution and mean teacher network. *Journal of Infrastructure Intelligence and Resilience*, *3*(4). doi: 10.1016/j.iintel.2024.100095

Zhang, X., Liu, L., Wu, F., & Wan, X. (2021). Digital Twin-Driven Surface Roughness Prediction Based on Multi-sensor Fusion. *Lecture Notes in Electrical Engineering*, *737*, 230–237. doi: 10.1007/978-981-33-6318-2_29

Zhang, Z., & Pan, W. (2020). Lift planning and optimization in construction: A thirty-year review. In Automation in Construction (Vol. 118). Elsevier B.V. doi: 10.1016/j.autcon.2020.103271

Zhang, Z., & Pan, W. (2021). Multi-criteria decision analysis for tower crane layout planning in high-rise modular integrated construction. *Automation in Construction*, *127*. doi: 10.1016/j.autcon.2021.103709

Zhao, R., Yan, D., Liu, Q., Leng, J., Wan, J., Chen, X., & Zhang, X. (2019). Digital twin-driven cyber-physical system for autonomously controlling of micro punching system. *IEEE Access*, *7*, 9459–9469. doi: 10.1109/ACCESS.2019.2891060

Zhao, Y., Lv, W., Xu, S., Wei, J., Wang, G., Dang, Q., Liu, Y., & Chen, J. (2023). *DETRs Beat YOLOs on Real-time Object Detection*. Retrieved from http://arxiv.org/abs/2304.08069

Zheng, B., Hussain, M., Yang, Y., Chan, A. P. C., & Chi, H. L. (2023). Trade-offs between accuracy and efficiency in BIM-LCA integration. *Engineering, Construction and Architectural Management*. doi: 10.1108/ECAM-03-2023-0270

Zheng, Z., Zhang, Z., & Pan, W. (2020). Virtual prototyping- and transfer learning-enabled module detection for modular integrated construction. *Automation in Construction*, *120*. doi: 10.1016/j.autcon.2020.103387

Zhou, C., Dai, F., Xiao, Z., & Liu, W. (2023). Location Optimization of Tower Cranes on High-Rise Modular Housing Projects. *Buildings*, *13*(1). doi: 10.3390/buildings13010115

Zhou, W., Zhao, T., Liu, W., & Tang, J. (2018a). Tower crane safety on construction sites: A complex sociotechnical system perspective. *Safety Science*, *109*, 95–108. doi: 10.1016/j.ssci.2018.05.001

Zhou, W., Zhao, T., Liu, W., & Tang, J. (2018b). Tower crane safety on construction sites: A complex sociotechnical system perspective. *Safety Science*, *109*, 95–108. doi: 10.1016/j.ssci.2018.05.001

Zhou, Yan. (2024). A YOLO-NL object detector for real-time detection. *Expert Systems with Applications*, *238*. doi: 10.1016/j.eswa.2023.122256

Zhou, Ying, Zhang, E., Guo, H., Fang, Y., & Li, H. (2021). Lifting path planning of mobile cranes based on an improved RRT algorithm. *Advanced Engineering Informatics*, *50*. doi: 10.1016/j.aei.2021.101376

Zhu, A., Zhang, Z., & Pan, W. (2022a). Crane-lift path planning for high-rise modular integrated construction through metaheuristic optimization and virtual prototyping. *Automation in Construction*, *141*. doi: 10.1016/j.autcon.2022.104434

Zhu, A., Zhang, Z., & Pan, W. (2022b). Crane-lift path planning for high-rise modular integrated construction through metaheuristic optimization and virtual prototyping. *Automation in Construction*, *141*. doi: 10.1016/j.autcon.2022.104434

Zhu, Z., Xi, X., Xu, X., & Cai, Y. (2021). Digital Twin-driven machining process for thin-walled part manufacturing. *Journal of Manufacturing Systems*, *59*, 453–466. doi: 10.1016/j.jmsy.2021.03.015

Zhuang, K., Shi, Z., Sun, Y., Gao, Z., & Wang, L. (2021). Digital twin-driven tool wear monitoring and predicting method for the turning process. *Symmetry*, *13*(8). doi: 10.3390/sym13081438

Zuo, Y., Wu, Y., Min, G., & Cui, L. (2019). Learning-based network path planning for traffic engineering. *Future Generation Computer Systems*, *92*, 59–67. doi: 10.1016/j.future.2018.09.043