Copyright Undertaking

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

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

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

IMPORTANT

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

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

SCANNING TRANSMISSION ELECTRON MICROSCOPY IMAGE ATOMIC POSITION DETECTION BY USING DEEP LEARNING

YAN ZHANGYUAN

MPhil

The Hong Kong Polytechnic University

2024

The Hong Kong Polytechnic University

Department of Applied Physics

Scanning Transmission Electron Microscopy Image Atomic Position Detection by Using Deep Learning

YAN Zhangyuan

A thesis submitted in partial fulfilment of the requirements

for the degree of Master of Philosophy

December 2023

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 ma-terial that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_______________________________________ (Signature)

YAN Zhangyuan (Name of candidate)

Abstract

Scanning Transmission Electron Microscopy (STEM) is a powerful imaging technique that offers high spatial resolution and detection sensitivity. However, traditional image processing techniques face limitations when dealing with complex and noisy STEM images. This paper explores the use of deep learning methods to analyze STEM images and proposes AtomID-net, an atomic finding model based on UNet. The model is trained and tested using real STEM data, addressing the limitations of simulated images. Furthermore, a new data labeling method is introduced for quick and accurate annotation of noisy images. The performance of AtomID-net is compared with traditional image processing techniques and other deep learning methods, demonstrating its superior performance and flexibility in detecting atomic positions. This research contributes to the advancement of deep learning in STEM image analysis and provides a reliable evaluation scheme.

Acknowledgements

I am grateful to **Dr. Zhao Jiong** for giving me the opportunity to be his student. His profound knowledge and research experience have been a great source of inspiration for me. His guidance extends beyond academia and includes practical skills and career development advice. His guidance and suggestions throughout the research process have allowed me to fully utilize my potential and achieve satisfactory research results.

I would like to express my gratitude to **Prof. Chen Ke**, my co-supervisor, for his assistance and guidance during the research process. His expertise and unique insights have provided valuable references and guidance for my research.

I would also like to thank **Dr. Zheng Xiaodong** for providing me with his valuable TEM (Transmission Electron Microscopy) data and for clarifying my doubts on TEM-related matters.

Similarly, I am grateful to **Dr. Wong Lok Wing** for sharing his TEM data with me and for discussing research experiences and problems.

I would like to express my appreciation to **Dr. Fu Qiang** for his advice and insights in assisting me with the challenges I encountered during my research.

I am deeply grateful to **Lin yanyu** for his guidance and valuable suggestions in my research project.

I am thankful to **Tsang Chi Shing Ben** for his assistance. His encouragement has made my learning and research journey smoother.

I would also like to thank **Shi Jie**, **Shen Wenqian** and other colleagues in the Department of Applied Physics and University for discussing, encouraging and helping each other.

Table of Contents

Certificate of Originality iii

Abstract iv

Acknowledgements v

Table of Contents vii

List of Figures ix

List of Tables x

List of Publications xi

Chapter 1. Introduction and Literature review 1

1.1. Scanning Transmission Electron Microscopy 1

1.2. Deep learning 4

1.3. Deep learning in STEM image 6

Chapter 2. Methodology 8

2.1. Architecture 8

2.2. Dataset 12

2.3. Evaluate method 19

Chapter 3. Results and Discussion 23

3.1. AtomID-net 23

3.2. Quality Assessment 27

3.3. Scale Assessment 30

3.4. Failure Case 33

3.5. Strain Mapping 35

3.6. Time Cost 37

Chapter 4. Conclusion 39

List of Figures

Figure 1. Overall schematic architecture of the AtomID-Net. 9

Figure 2. Simulation images example. 14

Figure 3. Training set of real HAADF images. 18

Figure 4. The annotation procedure for test set. 19

Figure 5. User interface of AtomID. 25

Figure 6. A few examples test with different methods. 30

Figure 7. Comparison with Findfoci in different magnification images. 32

Figure 8. AtomID model's results for HAADF images with different scale. 33

Figure 9. Detection on outliers. 34

Figure 10. Model performance for aperiodic atomic detection tasks. 35

Figure 11. Strain map calculation and analysis. 37

List of Tables

Table 1. The Chamfer Distance, Jaccard and F1 score of the test set. 26

Table 2. Quality Assessment of Different Methods. 31

Table 3. Time consume and result. 39

List of Publications

1. Lin, Y., **Z. Yan**, C. S. Tsang, L. W. Wong, X. Zheng, F. Zheng, J. Zhao & K. Chen (2023) A Multiscale Deep-Learning Model for Atom Identification from Low-Signal-to-Noise-Ratio Transmission Electron Microscopy Images. *Small Science*, 3, 2300031.

2. Wong, H. Y., L. W. Wong, C. S. Tsang, **Z. Yan**, X. Zhang, J. Zhao & T. H. Ly (2023) Superhydrophobic Surface Designing for Efficient Atmospheric Water Harvesting Aided by Intelligent Computer Vision. *ACS Applied Materials & Interfaces,* 15, 25849-25859.

Chapter 1. Introduction and Literature review

1.1. Scanning Transmission Electron Microscopy

Scanning Transmission Electron Microscopy (STEM) is a microscopy technique based on transmission electrons that obtains images by transmitting a high-energy electron beam through the sample and measuring the intensity and position of the transmitted electrons. [1-3] Compared to traditional Transmission Electron Microscopy, STEM offers higher spatial resolution and detection sensitivity. [3, 4] This is achieved by using a very small electron detector and focusing the electron beam to a very small size. STEM is widely applied in fields like materials science, nanotechnology, and biological sciences. [5-8]

The acquisition of STEM images involves two main steps: the generation of the transmission electron beam and the preparation of the sample. The transmission electron beam can be generated using a thermal emission electron gun or a field emission electron gun. In terms of sample preparation, thin sample specimens need to be prepared and placed under the transmission electron beam for observation. Furthermore, different directions of the sample can be obtained by rotating the sample or adjusting the incident angle of the electron beam to acquire more comprehensive sample information. [9-12]

With the recent development of aberration correctors for TEM, [13-15] the spatial

resolution of TEM (STEM) images can be as small as a few tens of picometers (pm). Electron microscopy data with atomic resolution contains a wealth of spatial distribution information about atoms and electron structures. [16, 17] Specifically, atomic structural information can be easily interpreted using scanning TEM (STEM) techniques. [18-21]

The first step in the structural analysis based on High-Resolution Transmission Electron Microscopy is to quantitatively associate local intensity peaks/valleys in the sample with atoms/atomic columns. [22] In the field of STEM image processing, many traditional image processing techniques have been widely applied. For example, traditional image filters such as mean filters, [23] median filters, [24] and Gaussian filters [25] can be used to remove noise from the images. Edge detection algorithms like Sobel [26] and Canny [27] can help identify edges and contours in the images. Additionally, Gaussian fitting algorithms [28, 29] and template matching algorithms [30] can be used to determine the positions of atomic columns in the images. Traditional feature extraction methods such as morphological operations, texture analysis, and shape descriptors can be used to extract features from the images. [31-34] These traditional image processing techniques provide a foundation for the processing and analysis of STEM images but may have limitations when dealing with complex and noisy STEM images. [35]

High-quality atomic-scale STEM images are typically valuable, requiring advanced TEM equipment and TEM operators with extensive knowledge in electron optics. Furthermore, environmental disturbances such as air flow, noise, vibrations,

AC/DC magnetic fields, and temperature fluctuations can introduce additional non-uniform noise into TEM images, significantly reducing their image quality. [36] Although the resolution and signal-to-noise ratio of TEM/STEM images can be improved by increasing the imaging electron dose, many materials, including two-dimensional materials, are beam-sensitive and prone to radiation damage. [37-40] Therefore, in practical TEM/STEM experiments, low-dose imaging conditions are typically required, which excludes the possibility of capturing high-quality images.

Furthermore, during the STEM imaging process, spontaneous adsorption of hydrocarbons on the sample surface in the TEM chamber can cause contamination. [41, 42] Due to the ultra-thin thickness of the sample and the resulting ultra-low contrast of the STEM signal, hydrocarbon contamination can manifest as super-bright/dark spots/regions, adding a noisy background that obscures the true sample information.

Traditional image processing techniques often face limitations when dealing with these non-ideal and low signal-to-noise ratio images. For example, when denoising low-quality images using Wiener filters, [43] although they can improve the signal-to-noise ratio to some extent, they are only suitable for periodic samples. Filtered atomic defects, dopants, surface, and other non-periodic structures may deviate from the original contrast intensity and atomic positions. Additionally, when facing low-quality images, traditional algorithms require laborious manual operations to extract valid features and information, and their accuracy is not always satisfactory. Furthermore, researchers often adjust and manipulate based on subjective perception in the absence

of quantitative performance metrics. Therefore, there is a need to establish an efficient and convenient method for automatic detection of atomic positions and chromatography, as well as a reliable evaluation scheme to address these issues.

1.2. Deep learning

Deep learning, as a branch of machine learning, has experienced remarkable success in recent years and has gained widespread attention in fields such as image recognition, speech recognition, and natural language processing. [44] It achieves this by constructing multi-layer neural networks that simulate the neural networks found in the human brain, enabling complex data to be learned and understood. [46, 47] The development of deep learning has been facilitated by advancements in computational power and the availability of large-scale datasets, providing a powerful tool to tackle complex real-world problems. [48]

Deep learning has undergone a lengthy development process over the past few decades. As early as the 1950s, artificial neural networks were proposed as a means to simulate the neural networks of the human brain. However, due to limitations in computational power and data scarcity, [49, 50] the progress of deep learning was somewhat restricted. It was only in recent years, with the rise in computational power and the availability of large-scale datasets, that deep learning has experienced rapid development and made significant breakthroughs in various domains.

At the core of deep learning lies the neural network, composed of multiple layers

of neurons. [51] Each neuron receives a set of inputs, performs a weighted sum, and generates an output through a non-linear activation function. [52] The hierarchical structure of the neural network allows for the transmission and transformation of information between different layers, facilitating abstract representations and feature extraction from the data. Commonly used models in deep learning include Convolutional Neural Networks (CNNs) [53-55] and Autoencoders [56]. CNNs are specifically designed for image processing in deep learning, employing convolutional layers [57] and pooling layers [58] to extract features and reduce dimensionality. Autoencoders, on the other hand, are unsupervised learning neural network models used for learning compressed representations of data.

Deep learning has found significant applications and made breakthroughs in numerous fields. In image recognition and classification tasks, it has achieved notable success in areas such as face recognition, object detection, and image captioning. In the domain of speech recognition and synthesis, deep learning is widely applied to tasks such as speech recognition, speech synthesis, and speech emotion analysis. In natural language processing, deep learning has made important strides in tasks like text classification, machine translation, and sentiment analysis. [59-63] Furthermore, deep learning plays a vital role in fields such as medical diagnosis, financial forecasting, and autonomous driving.

1.3. Deep learning in STEM image

Due to the powerful information extraction capabilities of deep learning networks and the enormous potential demonstrated in complex systems, researchers have begun exploring their possible roles and value in more domains. Recently, some researchers have employed deep learning methods to analyze electron microscopy images, attempting to leverage their strong feature extraction abilities to identify microstructures and deepen the understanding of the correlation between microstructures and material properties.

For example, Ziatdinov et al. [64, 65] proposed an efficient atomic position finding tool to identify atomic positions by combining deep learning networks with image processing algorithms. They simulated lattices containing artificially created defects using simulation software and trained an FCN network [66] with the simulated theoretical images. The trained network outputs a confidence map that provides a probability for each pixel to differentiate whether it belongs to an atomic column or the background. Laplacian of Gaussian [67] spot detection is used to extract the centers of circular atoms from the confidence map. Madsen et al. [68] trained a CNN network for local recognition of TEM images by utilizing a large number of simulated images. Testing this method on gold nanoparticles on monolayer defect graphene and cerium oxide substrates, they found that the network has the ability and potential to detect atomic column heights based on image contrast. Lin et al. [69] developed a forward model for simulating images by combining real scans with Poisson noise, making the simulated images closer to real

images. Using this simulated image model, they trained a U-Net [70] network called atomsegnet for atomic segmentation and detection, denoising, background removal, and super-resolution processing tasks.

Deep learning has demonstrated powerful capabilities in STEM image processing, but this field also faces significant limitations in terms of datasets. [71, 72] Unlike datasets in other domains, for training sets of real scans, researchers need to annotate atomic positions in STEM images. However, this process is costly, and manually annotating these data is virtually impossible when faced with millions of atoms in hundreds or thousands of images. Additionally, even experienced researchers sometimes struggle to determine atomic positions in noisy images. Therefore, in current deep learning methods for analyzing STEM images, researchers have resorted to using simulated images as a substitute for real images to reduce annotation costs. However, simulated images represent ideal theoretical images and cannot capture the rich structures, types of defects, and noise levels present in real images, which directly impacts the performance of the models.

To address these issues, this paper proposes AtomID-net, an atomic finding model based on U-Net. The model is trained and tested using real STEM data. By comparing this model with other traditional image processing techniques and deep learning methods, we find that AtomID outperforms these methods and does not require fixed-sized atoms. Additionally, we introduce a new data labeling method that enables us to annotate noisy images quickly and accurately.

Chapter 2. Methodology

This chapter mainly introduces the framework of deep learning model, the creation of training set, the training process, and the evaluation method. Part of the information in **Chapter 1** relates to this chapter. We will cover the methods we use in as much detail as possible.

2.1. Architecture

2.1.1. Overall schematic architecture

The AtomID-Net architecture used in this study is based on the popular U-Net encoder-decoder structure, originally designed for image segmentation. In this paper, the U-Net architecture is adapted for the task of atomic segmentation. The model takes in TEM images as inputs and their corresponding semantic labels as targets to train the model. The overall schematic architecture is shown in **Figure 1**.

The AtomID-Net architecture consists of convolutional blocks with 3x3-kernel and proper skip links. Each convolutional block has two convolutional layers and two rectified linear units (ReLU) [73]. Skip links [74] map outputs from the blocks of the encoder to the corresponding blocks of the decoder. Instead of the original linear layers, a convolutional block is used at the bottleneck, and an extra branch is designed to predict the global average atomic spacing.

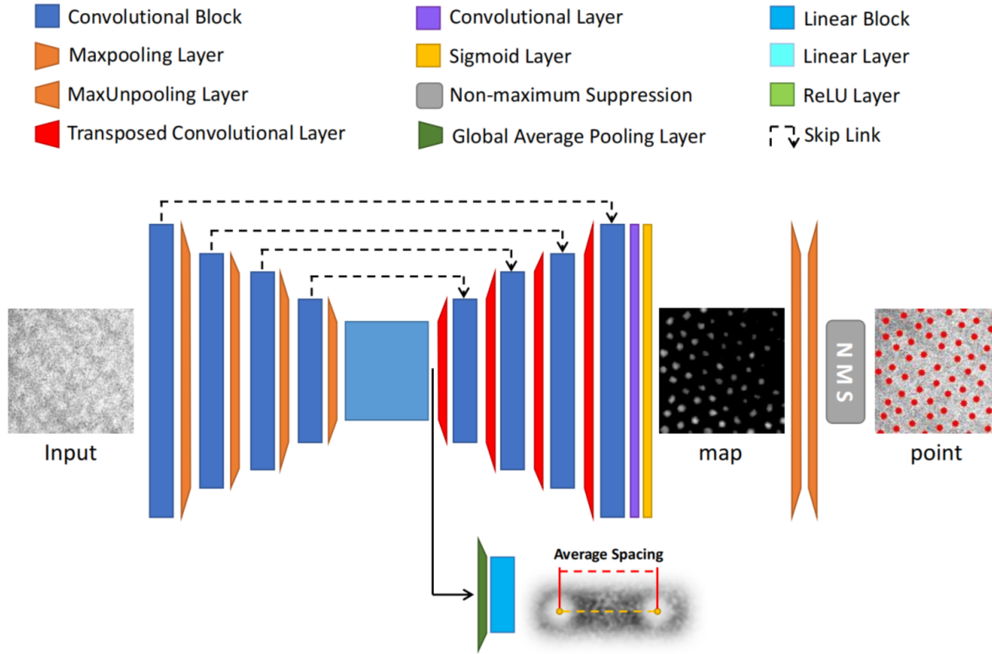

Figure 1. Overall schematic architecture of the AtomID-Net. The main branch outputs the confidence maps of atomic columns.

2.1.2. Loss Function

The main branch for segmentation is parameterized by θ and denoted by $f_\theta(\cdot)$, while the extra branch for spacing prediction is parameterized by μ and denoted by $g_\mu(\cdot)$. Although θ and μ partially overlap in the encoder, the side-effect of the partial merge is ignored in training.

The training set consists of input image set $\{\boldsymbol{X}_i\}_{i=1}^N$ and the corresponding target image set $\{\boldsymbol{Y}_i\}_{i=1}^N$. The targets are images with annotations at the foci of the atoms from the inputs. The i-th target image contains n_i annotated centers. An assumption is made that every observable atom has a shape of an approximate disc, and its geometric center gives a local maxima of intensity. All locations of atoms in the i-th target image can be

constructed as:

$$\boldsymbol{Y_i} = \sum_{\boldsymbol{y}\in\mathbb{R}^2} \delta(\boldsymbol{y}) \quad \textbf{Eq. 1}$$

where $\boldsymbol{Y_i}$ denotes the i-th dot-annotated label image, and y is the coordinate vector. The $\delta(\boldsymbol{y})$ is a unit impulse function. Then we apply scaled Gaussian filters on all the n_i centers, it can be conducted in a convolution operation:

$$\widetilde{\boldsymbol{Y}}_{\boldsymbol{i}} = \boldsymbol{Y_i} \otimes \boldsymbol{e}^{-\frac{y^T y}{2\sigma_i^2}} \quad \textbf{Eq. 2}$$

$$\sigma_i \propto \bar{s}_i \quad \textbf{Eq. 3}$$

$$\bar{s}_i = \frac{1}{n_i} \sum^{n_i} \min_{j \neq k} \left\| \boldsymbol{y}_j - \boldsymbol{y}_k \right\| \quad \textbf{Eq. 4}$$

where the global spacing $\bar{s}_i$ is the mean of all distances of all atoms to their closest counterparts in the i-th sample and the parameter σ_i is a variance factor proportional to $\bar{s}_i$. The $\otimes$ denotes a 2D convolution operation, with the kernels it smooths the labels by turning hollow images with sparse dots into confidence maps. The $\widetilde{\boldsymbol{Y}}_{\boldsymbol{i}}$ represents the i-th smoothed confidence map. The coefficient of the Gaussian kernel is dropped to make a scaled version of the Gaussian filter. It avoids fluctuations of intensity with the spacing varies and enables subsequent operations to favor the centers equally likely.

The task of the main branch is to segment atoms from the input images and output confidence maps that are close to the given target maps. Each output pixel's intensity

represents the possibility of being an atomic center. The higher intensity, the closer to the nearest center. The optimal target can be written as:

$$\min_{\theta} \frac{1}{2N} \sum_{i=1}^{N} \left\| f_{\theta}(\boldsymbol{X}_i) - \widetilde{\boldsymbol{Y}}_i \right\|_2^2 \qquad \textbf{Eq. 5}$$

where the X is the input of network f_{θ} with learnable parameters θ, $\tilde{Y}_i$ is the smoothed target confidence map (generated by (2)). The second task is to predict global spacing $\bar{s}_i$through another branch g_{μ}. The optimal target can similarly be written as:

$$\min_{\mu} \frac{1}{2N} \sum_{i=1}^{N} \left\| g_{\mu}(\boldsymbol{X}_i) - \bar{s}_i \right\|_2^2 \qquad \textbf{Eq. 6}$$

The two target functions are combined with a tradeoff constant to balance training of both terms, resulting in total loss function L:

$$L = \frac{1}{2N} \sum_{i=1}^{N} (\lambda \left\| f_{\theta}(\boldsymbol{X}_i) - \widetilde{\boldsymbol{Y}}_i \right\|_2^2 + \left\| g_{\mu}(\boldsymbol{X}_i) - \bar{s}_i \right\|_2^2) \qquad \textbf{Eq. 7}$$

2.1.3. Non-maximum suppression

During the detection stage, the Maxpooling and Maxunpooling layers downsample a predicted confidence map and then upsample it back to reduce redundant local maximas and noises. The remaining local maximas are considered as detection proposals, and the predicted atomic spacing as the side lengths of detection bounding boxes. These proposals and spacings are sent to the final detection module, non-maximum suppression (NMS). [75]

NMS is a widely used technique for object detection. It assesses the similarity

between bounding boxes by measuring their overlap, usually with the Intersection over Union (IoU) metric. Bounding boxes with an IoU above a certain threshold are considered to overlap. In the NMS process, the box with the highest confidence score is initially chosen as the output, and then, any remaining boxes with an overlap greater than the threshold are progressively eliminated. This iterative process identifies high-confidence bounding boxes that do not overlap until all boxes have been processed.

The NMS algorithm finds the most possible atomic centers among the proposals by filtering out those which fail the IOU restrictions and outputs the coordinates. Comparing to other matching algorithms, the time consumption is resolution-irrelevant when the count of filtered proposals is fixed. The architecture of the AtomID-Net model is displayed in Figure 1.

2.1.4. Training parameters

The AtomID-Net model was trained at an initial learning rate of 10^{-3} on a single GPU 1080Ti for 200 epochs. The learning rate decay to 0.1 times each 50 epochs. λ in loss function is set to 0.1. The IOU allow rate in the NMS algorithm is set to 0.3.

2.2. Dataset

The data set consists of two parts, the real image training set, and the simulated image training set.

2.2.1. Simulated Image Dataset

ABTEM [76, 77] is a tool used to simulate scanning transmission electron microscopy (STEM) images. It is a computer program based on the principles of transmission electron microscopy (TEM) theory and electron optics simulations. By simulating the interaction between electrons and samples, including scattering and transmission processes, ABTEM generates realistic STEM images. It considers the propagation of electron beams, sample structures and compositions, as well as detector responses. Users can input crystal structures, defect information, lattice parameters, and select various imaging parameters, such as electron beam energy and incident angles. ABTEM provides an open-source Python code library that we favor for its ability to simulate different imaging modes and its fast simulation speed.

We utilized ABTEM to simulate synthetic images in the creation of our simulated image dataset. We simulated the structures of two materials, 1T” ReS_2 and 2H MoS_2, using their structural data. To enhance the realism of the simulated images, we intentionally introduced disruptive elements such as vacancy defects, voids, and doped atoms to disrupt the periodic structure during the simulation process. Additionally, we added noise sources like Poisson noise, Gaussian noise, and local bright spots to the images as shown in **Figure 2**. The atomic coordinates are recorded in the ground truth images in the form of Gaussian spots.

(a)

(b)

(c)

(d)

Figure 2. Simulation images example. When simulating the image, we added and changed noise parameters such as **(a)** defocus **(b)** astigmatism **(c)** Poisson noise and **(d)** Gaussian noise.

These simulated images will be used to train four shallow neural networks, which will be used to label other noise images. The four networks were trained from thousands of simulations of 2H phase with low magnification, 2H phase with high magnification, 1T" phase with low magnification, and 1T" phase with high magnification, respectively,

and were only used to test HAADF images with the same input structure.

2.2.2. Real Image Dataset

The real image training dataset comprises two types of materials: MoS_2 and ReS_2. These materials are grown as monolayer films using the Chemical Vapor Deposition (CVD) technique. Subsequently, the films are transferred onto Cu quantifoil grids and observed using a Spectra 300 microscope.

CVD is a technique commonly used to grow thin films or monolayers of materials. It involves the deposition of vaporized precursor chemicals onto a substrate, where they react and form the desired material. In the case of MoS_2 and ReS_2, both materials were prepared using atmospheric pressure CVD.

For the growth of monolayer MoS_2, a (001)-cut fluorophlogopite mica substrate was used. The mica was placed face-down above a quartz boat in a 2-inch CVD furnace. The quartz boat contained 3 mg of molybdenum trioxide (MoO_3) as the molybdenum precursor. Another crucible placed upstream contained 120 mg of sulfur (S) as the sulfur precursor.

For the growth of monolayer ReS_2, a single-side polished c-face sapphire substrate was used. In the upstream zone, sulfur powder was placed as the sulfur precursor, while in the downstream zone, ammonium perrhenate (NH_4ReO_4) was used as the rhenium precursor.

In both cases, the CVD process allowed the precursor chemicals to react and form monolayer films of MoS_2 and ReS_2 on the respective substrates. The specific growth recipes, including gas flow, temperature, and heating duration, were optimized to achieve the desired material quality and thickness.

The two materials were provided by Dr. Jiong Zhao and his group.

TEM sample preparation. To transfer the as-grown MoS_2 onto Cu quantifoil grid, the MoS_2 was coated with a layer of polymeththyl methacrylate (PMMA) (A4) by spin-coating of 3000 rpm for 1 min, followed by baking at 100°C for 5min. The process was repeated for two layers of PMMA coating. Next, the PMMA/MoS_2 on Mica was immersed in deionized (DI) water for 5 min and then detached by gentle pulling. Subsequently, we fished out the floating PMMA/MoS_2 film with a grid and placed the grid on a hot plate at 60°C for 10 min, stored under vacuum for 12 hours. The PMMA was then removed by rinsing acetone, IPA, and DI water. After that, the grid was annealed at 200°C for 1100 min to remove residues of hydrocarbons on the surface.

TEM experiments. Experimental HAADF-STEM imaging was performed using an aberration-corrected TEM (Spectra 300 TEM) operated at 300 kV, and the electron beam current was around 10~20 pA. The CL aperture during image capture was 115 mm and the collection angle was 45 to 180 mrad, the dwell time was set to 10 μs and the original images were 1024 × 1024 pixels.

After performing TEM experiments, we obtained the real image training dataset.

This dataset comprises 162 images with a size of 1024×1024 and 50 images with a size of 800×800 for training and testing purposes. The dataset primarily consists of 1L ReS_2 in the 1T" phase and MoS_2 in the 2H phase. Due to the high cost and inevitable ambiguity of manually annotating real images containing thousands of atoms, we employed the Findfoci algorithm, and a shallow neural network trained on simulated images to automatically label the training set. The atomic coordinates are recorded in the ground truth images as Gaussian spots. To introduce diversity into the training data, we applied random adjustments such as resizing, cropping, and adding noise to the training images, resulting in a large and diverse dataset. **Figure 3** shows the examples of training set.

The image data of the training set was provided by Dr. Zhao Jiong and his group.

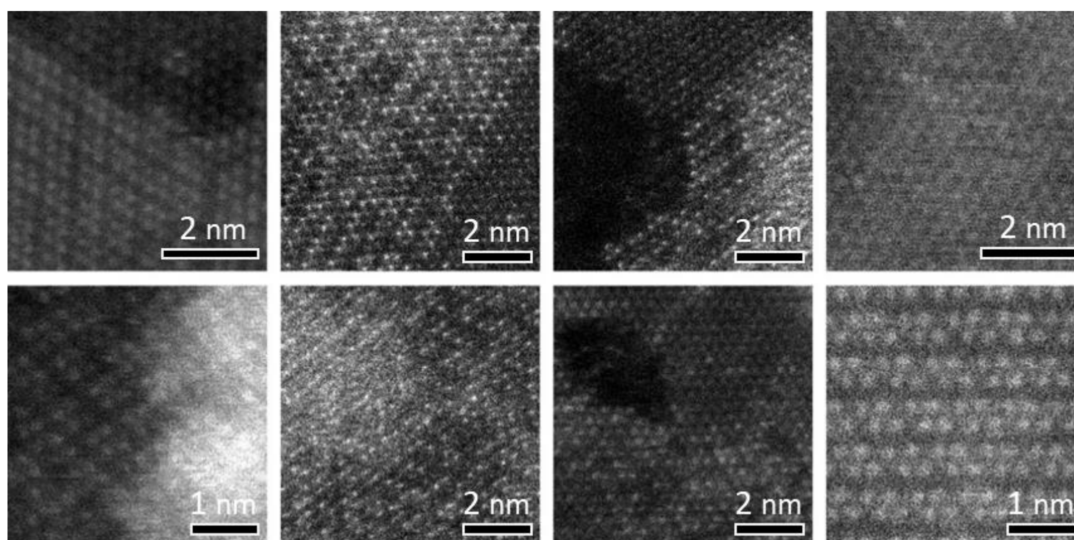

Figure 3. Training set of real HAADF images.

Findfoci is a peak-finding algorithm used for localizing atomic points. The algorithm identifies points in an image that have higher intensity compared to their surrounding points and gradually expands these peak regions, assigning lower-intensity points to the peak points within the regions. By adjusting the parameters, redundant or erroneous peaks can be merged, thereby improving performance on noisy images. This algorithm is suitable for batch processing of images and utilizes labeled images to train the algorithm for selecting the optimal parameter combinations.

The test dataset was generated using a different methodology. When capturing images with the Spectra 300 microscope, a specific protocol was followed: each real-world sample was imaged 10 times in a fixed region using STEM, with random variations in STEM current, defocus, and other parameters to capture images under different conditions. During this process, drift and spatial deviations inevitably occurred, leading to variations in atomic positions at different points. To mitigate drift, a cross correlation method [78, 79] was employed to align the images within each set. To refine the labels for each set, the best-visible image was automatically annotated as a reference using the Findfoci algorithm. [80] Subsequently, a two-dimensional Gaussian fitting method [25] was applied to fine-tune the reference and generate accurate labels. Visual duplicates were removed from the 30 sets, resulting in a final test dataset of 50 images. Specifically, sulfur atoms, which have a less dominant presence in the sample and are challenging to observe, were disregarded during the annotation process. The annotation process for the test dataset is illustrated in **Figure 4**. This annotation method

ensures the reliability of data from blurry images that are difficult for the human eye to discern.

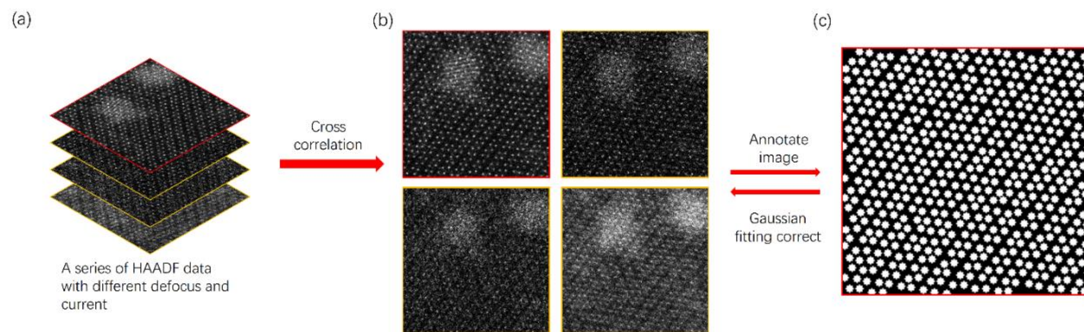

Figure 4. The annotation procedure for test set. **(a)** A group of HAADF images with different defocus and current. **(b)** The cross-correlation method corrected the image drift. **(c)** The image with the best quality was selected (the image with the red border) and annotated by Findfoci as a reference for the group, and the 2D Gaussian fitting is used to refine the reference into precise labels. (See text for more details)

2.3. Evaluation method

To evaluate the performance of the model and validate its training results, it is essential to have quantitative and reliable metrics. In this section, we will provide a comprehensive overview of three distinct metrics used for evaluating the model's performance.

2.3.1. Chamfer Distance

Chamfer Distance, [81] also known as Chamfer Matching or Chamfer Loss, is a distance metric used to measure the similarity or dissimilarity between two sets of points or shapes. It is commonly employed in computer vision and image processing tasks, such as object recognition, shape matching, and image segmentation.

The basic idea behind Chamfer Distance is to calculate the minimum distance between a set of reference points and a target set of points. It quantifies the dissimilarity by measuring how well the target points match the reference points. The name "Chamfer" comes from the French word for "bevel," as the distance calculation involves computing the distances along the beveled edges or contours of the shapes.

It describes the average distance between the predicted set and the ground truth. The algorithm is given by:

$$CD = \frac{1}{m_p} \sum_{i=1}^{m_p} \min_{j} \left\| C_i^{(p)} - C_j^{(t)} \right\| + \frac{1}{m_t} \sum_{j=1}^{m_t} \min_{i} \left\| C_i^{(p)} - C_j^{(t)} \right\| \qquad \textbf{Eq. 8}$$

Where m_p and m_t denote the total amount of predicted atoms and the one of the ground truth. $C_i^{(p)}$ and $C_j^{(t)}$ denotes the coordinates of the i-th center from the prediction and the j-th center from the ground truth.

2.3.2. Jaccard Score

Jaccard Score, [82] also known as Jaccard Index or Intersection over Union (IoU), is a statistical metric used to evaluate the similarity or dissimilarity between two sets. It is commonly employed in the field of data science, machine learning, and image segmentation to measure the accuracy of predictions or the overlap between sets.

The Jaccard Score is calculated as the ratio of the intersection of two sets to the union of those sets. It quantifies the extent of overlap between the sets, providing a

value between 0 and 1, where 0 indicates no overlap, and 1 indicates a perfect match.

It is given by:

$$Jaccard = \frac{Matches}{Matches + Unmatched^{(p)} + Unmatched^{(t)}} \quad \textbf{Eq. 9}$$

Matches denotes the true positive prediction, the number of matched points between the predicted and ground true sets. $Unmatched^{(p)}$ and $Unmatched^{(t)}$ represent the false positive predictions and the true negative targets respectively, the count of points from each set that fail to pair up.

2.3.3. F1 Score

F1 Score [83] is a statistical metric commonly used to evaluate the performance of binary classification models. It combines precision and recall into a single value, providing a balanced measure of the model's accuracy.

In binary classification, the F1 Score considers two classes: positive (or "true") and negative (or "false"). The goal is to classify instances correctly into one of these classes. The F1 Score is calculated based on the following two metrics:

Precision: Precision measures the proportion of correctly predicted positive instances among all instances predicted as positive. It quantifies the model's ability to avoid false positives. Precision is calculated as:

$$precision = \frac{Matches}{Matches + Unmatched^{(t)}} \quad \textbf{Eq. 10}$$

Recall: Recall, also known as sensitivity or true positive rate, measures the proportion of correctly predicted positive instances among all actual positive instances. It quantifies the model's ability to avoid false negatives. Recall is calculated as:

$$\boldsymbol{recall} = \frac{\boldsymbol{Matches}}{\boldsymbol{Matches + Unmatched^{(p)}}} \qquad \textbf{Eq. 11}$$

The F1 Score is the harmonic mean of precision and recall, providing a single measure that balances both metrics. It is calculated as:

$$\boldsymbol{F_1 = 2 * \frac{precision * recall}{precision + recall}} \qquad \textbf{Eq. 12}$$

The F1 Score ranges from 0 to 1, where 1 represents the best possible performance, indicating perfect precision and recall, while 0 indicates the worst performance, with no correct positive predictions.

The F1 Score is particularly useful when the dataset is imbalanced and the positive and negative classes have different frequencies. It considers both false positives and false negatives, making it a robust evaluation metric in such scenarios.

Chapter 3. Results and Discussion

In this chapter, we will test and evaluate the trained model and compare it with other atomic identification methods. We will discuss the advantages and disadvantages of this model in detail.

3.1. AtomID-net

We obtained the final AtomID-net model by training it with 162 real STEM images of size 1024 x 1024. To facilitate testing on images, an interactive GUI interface was designed, as shown in the **Figure 5**. Through this interface, two adjustable parameters in the model can be tuned to improve its performance. The two adjustable parameters are "Filter Lower Bound" and "Filter Window Size." The former adjusts the threshold to control the output of atomic positions and raising it appropriately can avoid atomic misidentification caused by noise or contamination. The latter is an effective means to control the sparsity of the output atoms, and by adjusting it, the model can be applied to a larger scale range.

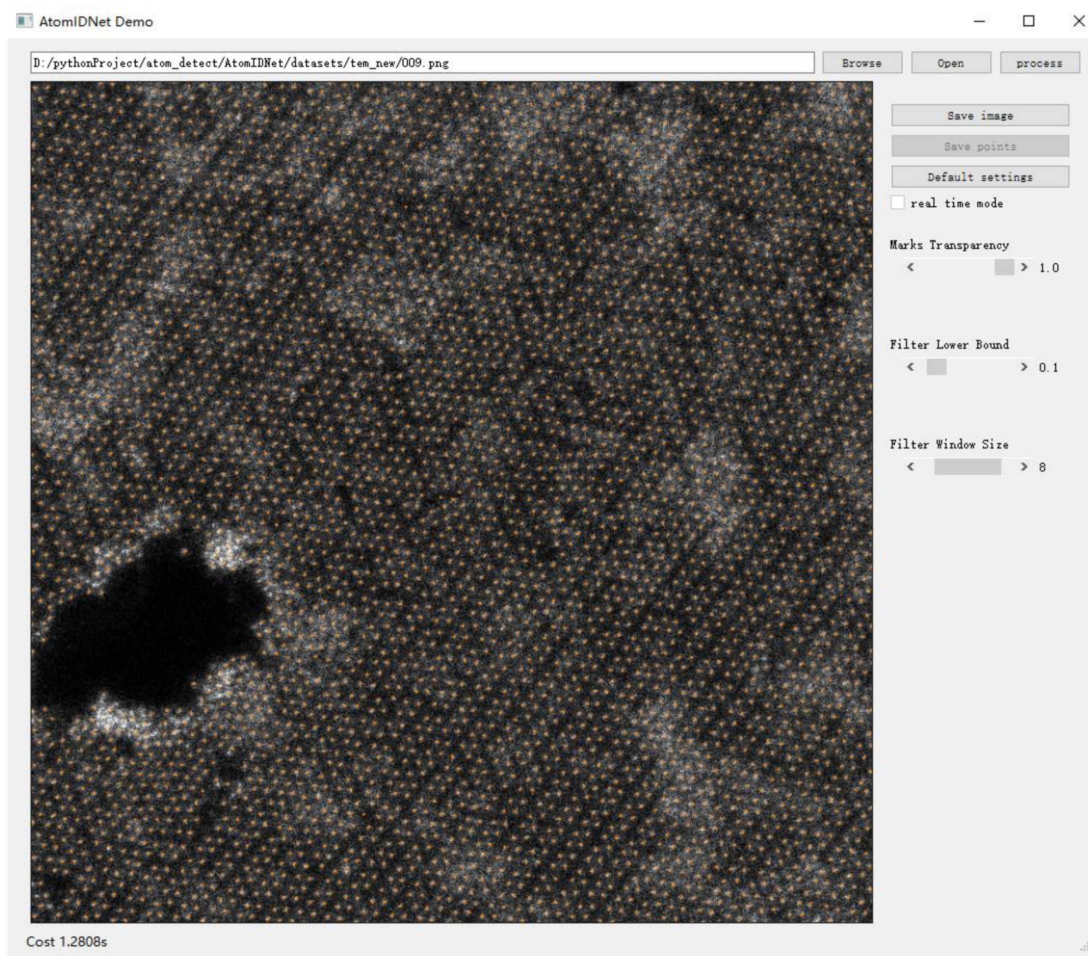

Figure 5. User interface of AtomID.

After setting the parameters, clicking "process" applies the trained model for testing, and the identified atomic coordinates are marked on the image as red dots. The "Mask Transparency" controls the transparency of the red dot annotations, allowing for better comparison and examination between the original image and the results. Additionally, if the "real-time model" option is selected, the model will provide real-time results based on the adjusted parameters without the need to click "process" again, which is advantageous for finding better parameter combinations.

By testing the model on our set of 50 test images of size 800 x 800, we obtained the Chamfer Distance (CD), Jaccard, and F1 scores. The results demonstrate that the AtomID model performs accurately in atomic recognition on this dataset, with an average F1 score of 0.964. The results are presented in the **Table 1**.

Table 1. The Chamfer Distance, Jaccard and F1 score of the test set.

Test no.	Material	CD	Jaccard	F1
1	ReS_2	1.974076	0.969437	0.984482
2	ReS_2	2.150585	0.961848	0.980553
3	ReS_2	1.881161	0.966359	0.982892
4	ReS_2	1.853786	0.976552	0.988137
5	ReS_2	1.883273	0.969621	0.984576
6	ReS_2	1.734714	0.97366	0.986654
7	MoS_2	1.84006	0.965609	0.982504
8	MoS_2	2.044782	0.953225	0.976053
9	MoS_2	2.284605	0.931407	0.964486
10	MoS_2	4.88553	0.606143	0.754781
11	MoS_2	5.305907	0.581759	0.935585
12	MoS_2	5.920004	0.489061	0.856872
13	MoS_2	1.775613	0.942716	0.970513
14	MoS_2	2.533263	0.88254	0.937605
15	MoS_2	1.944994	0.936378	0.967144
16	MoS_2	2.015008	0.923669	0.96032
17	MoS_2	1.83408	0.977197	0.988467
18	MoS_2	2.026449	0.9635	0.981411
19	MoS_2	1.915053	0.946586	0.97256

20	MoS_2	2.107459	0.9411	0.969656
21	MoS_2	1.628142	0.973293	0.986466
22	MoS_2	1.701772	0.964545	0.981953
23	MoS_2	1.955338	0.948896	0.973778
24	ReS_2	1.742736	0.929589	0.96351
25	ReS_2	1.842315	0.921191	0.958979
26	ReS_2	1.763286	0.929995	0.963728
27	ReS_2	2.245398	0.95531	0.977144
28	ReS_2	2.598618	0.938578	0.968316
29	ReS_2	5.939665	0.597012	0.747661
30	ReS_2	1.751999	0.988122	0.994026
31	ReS_2	1.280296	0.984279	0.992077
32	ReS_2	4.845748	0.621665	0.866699
33	ReS_2	1.568182	0.985171	0.99253
34	ReS_2	3.378481	0.842155	0.914315
35	ReS_2	1.394043	0.982124	0.990981
36	ReS_2	2.09328	0.94612	0.972314
37	ReS_2	1.898353	0.986462	0.993185
38	ReS_2	1.880469	0.987666	0.993795
39	ReS_2	1.232682	0.984906	0.992396
40	ReS_2	1.253336	0.986449	0.993178
41	ReS_2	1.325335	0.981266	0.990545
42	ReS_2	1.421264	0.977877	0.988815
43	ReS_2	1.346379	0.98343	0.991646
44	ReS_2	1.377934	0.980895	0.990355
45	ReS_2	1.31586	0.99105	0.995505

46	ReS_2	1.377934	0.980895	0.990355
47	ReS_2	1.623968	0.99014	0.995045
48	ReS_2	1.778877	0.9922	0.996085
49	ReS_2	2.453387	0.94574	0.972113
50	ReS_2	2.107842	0.953016	0.975943

3.2. Quality Assessment

To evaluate the performance of our model, we compared our model with three other methods, Findfoci, Atomap and AtomSegNet.

Findfoci [80] is a peak-finding algorithm used for localizing atomic points. To filter out background noise, we set the absolute background parameter within the range of 0.1-0.3. Subsequently, the image is subjected to Gaussian blurring to reduce spurious peaks and achieve image smoothing. The value for Gaussian blur is set between 0.8 and 2. We set the minimum size of the peaks to 10, but we adjusted this parameter for images with different magnifications to prevent atomic peaks from splitting into smaller ones.

Atomap [84] is a method used for locating atomic columns by finding aligned atomic columns and applying a two-dimensional elliptical Gaussian fit to each atomic column. The primary objective of this method is to identify the positions of atomic columns in an image and normalize them structurally, thereby improving the accuracy and robustness of the localization. By adjusting the pixel separation value, Atomap can adapt to images with different structures.

In Atomap, we utilize the peak-finding algorithm from the scikit-image Python package, which requires a minimum peak separation. To find the optimal pixel separation, we use the Atomap function *get_feature_separation()*. For our test set, the separation range is set between 4 and 10, varying depending on the magnification of the image.

A previous work named AtomSegNet [69] was proposed to augment synthetic training images by adding background and Poisson noises, so it can improve robustness against noises. AtomSegNet offers 15 different models for atomic position detection. After testing on our validation set, we found that the model *denoise&gremoval&superres* performed the best. Therefore, we mostly used this model to test the images in our validation set, but we also used the model *gaussianMask+* sometimes. We use the parameter "Up sample by 2" to enlarge the image by two times, which helps the model detect images with dense arrays of atoms. We sometimes use this parameter for low magnification images.

In the test of 50 STEM images, the average F1-score of the AtomID-Net model is 0.964, which exceed the other three methods. The results are shown in **Table 2**. Note that, all methods perform well for clear images, but our model performs better for low-quality images. **Figure 6** shows performance of AtomID-Net on some low SNR images.

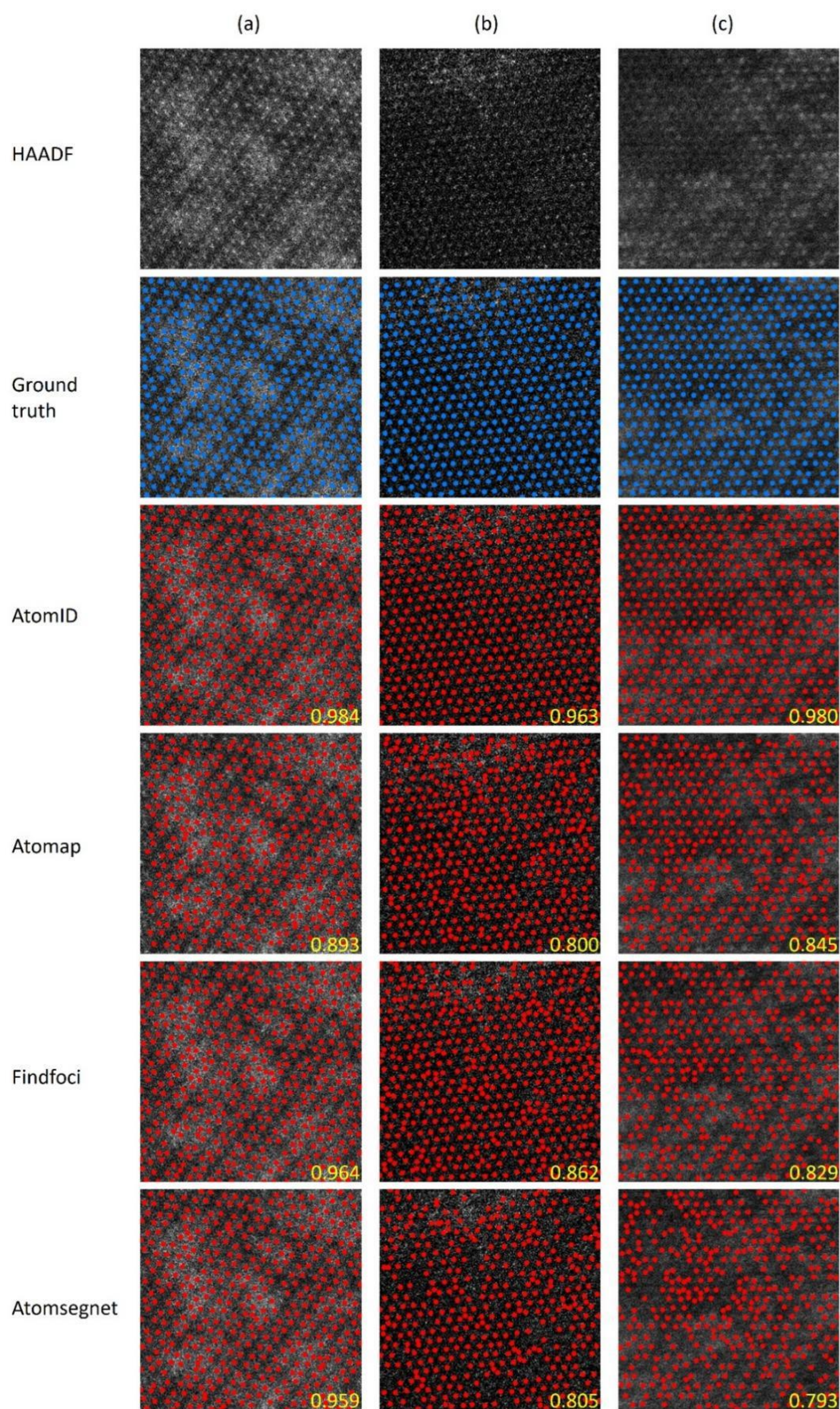

Figure 6. A few examples test with different methods. The F1 score is shown in yellow. Atomic scale: (a) 6.35nm × 6.35nm; (b) 6.35nm × 6.35nm; (c) 7.03nm × 7.03nm.

Table 2. Quality Assessment of Different Methods.

	AtomID	Atomap	Findfoci	AtomSegNet
CD	2.373	4.174	2.636	4.550
Jaccard	0.967	0.774	0.814	0.733
F1	0.964	0.864	0.891	0.832

3.3. Scale Assessment

To test the performance of AtomID-net at different scales, we used a series of HAADF data with different magnification to do test, and we compared our model with Findfoci, a peak seeking algorithm which we used to generate auto-annotated training samples. In comparison, we used HAADF images with different magnifications of the same structure and applied the same set of parameters to the Findfoci method and our model. The results in **Figure 7** show that, when testing on inconsistent scales across images, the Findfoci searches exhaustively for optimal hyper-parameters to achieve satisfactory results, while it can be addressed faster with our proposed model which make it more practical.

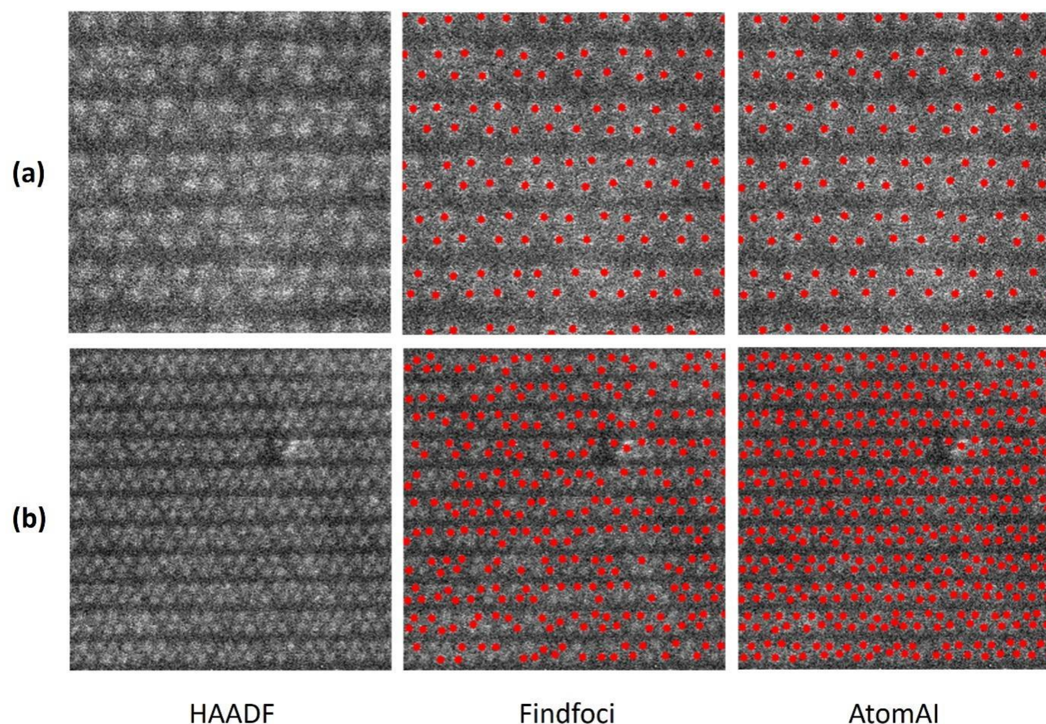

Figure 7. Comparison with Findfoci in different magnification images. Atomic scale: **(a)** 3.75nm × 3.75nm. **(b)** 7.03nm × 7.03nm.

Furthermore, the performance of AtomID in identifying atomic column positions is remarkable within the atomic scale range of 9nm to 36nm, as shown in the **Figure 8**. Without adjusting the parameters, this model achieves an F1 score of approximately 0.95, and there is no decrease in the score as the atomic scale range increases. This indicates that the AtomID model exhibits a certain level of robustness to variations in the atomic scale, enabling it to adaptively identify atomic positions in STEM images with different scales and magnifications while maintaining a high level of accuracy.

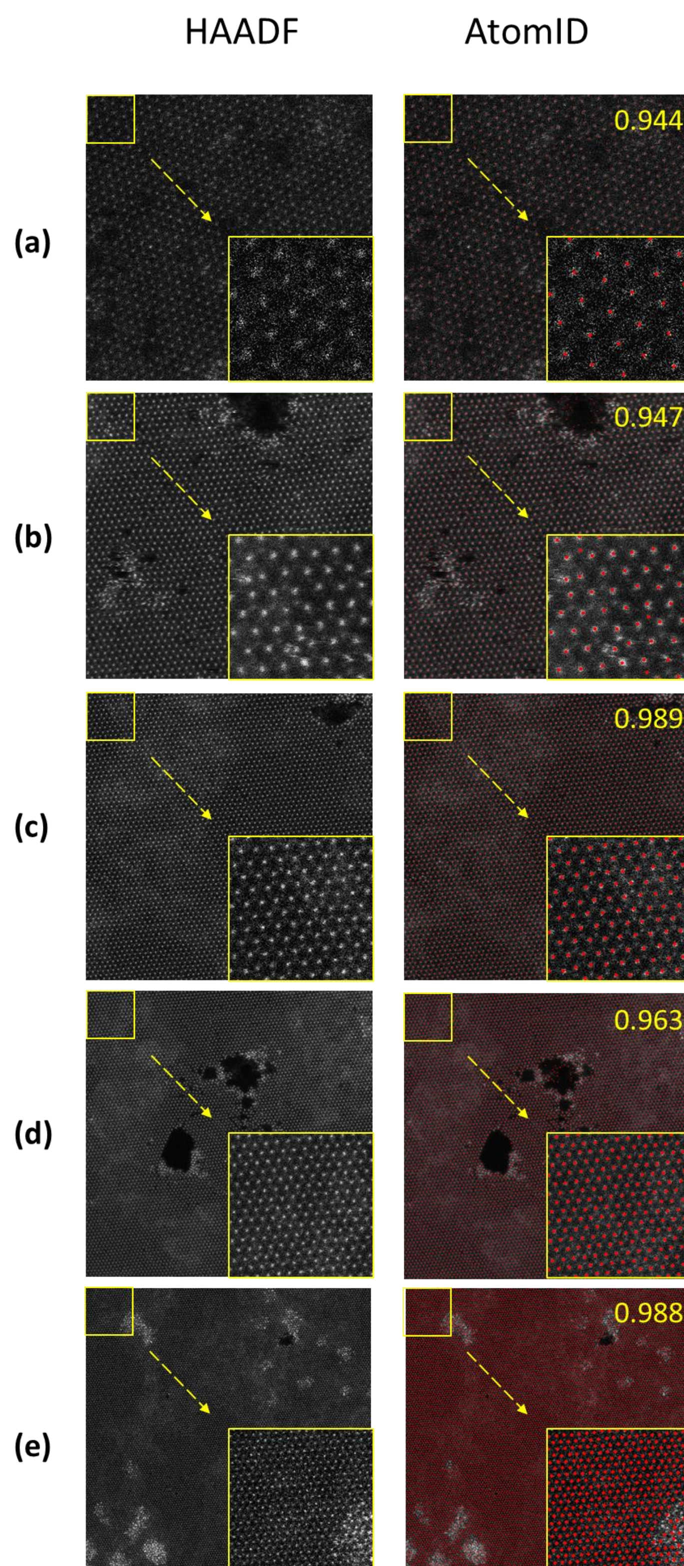

Figure 8. AtomID model's results for HAADF images with different scale. This model is flexible and is not limited to a fixed scale of atoms. F1-score is shown in yellow. Atomic scale: **(a)** 8.99nm × 8.99nm; **(b)** 12.7nm × 12.7nm; **(c)**18.0nm × 18.0nm; **(d)** 25.4nm × 25.4nm; **(e)** 36.0nm × 36.0nm

3.4. Failure Case

We also ran training and testing on outliers with other structures or nonperiodic materials. We concluded that we need prior assumptions to deal with low-SNR samples. A model trained on data only $MoS_2$of phase $2H$ has dramatic bias, when running test on with hexagonal structures (bilayer MoS_2), it tends to give a detection result satisfying phase $2H$ as shown in **Figure 9**. That is because in our dataset consisting of monolayer $1T''$ phase ReS_2 and $2H$ phase MoS_2, the sulfur atoms are mostly hidden in noisy backgrounds and has been ignored. In the structure of hexagon, atoms in the ignored positions are dominant and are expected to be located. Consequently, training with multiple incompatible structures in a shared pipeline may cause conflicts and performance reduction.

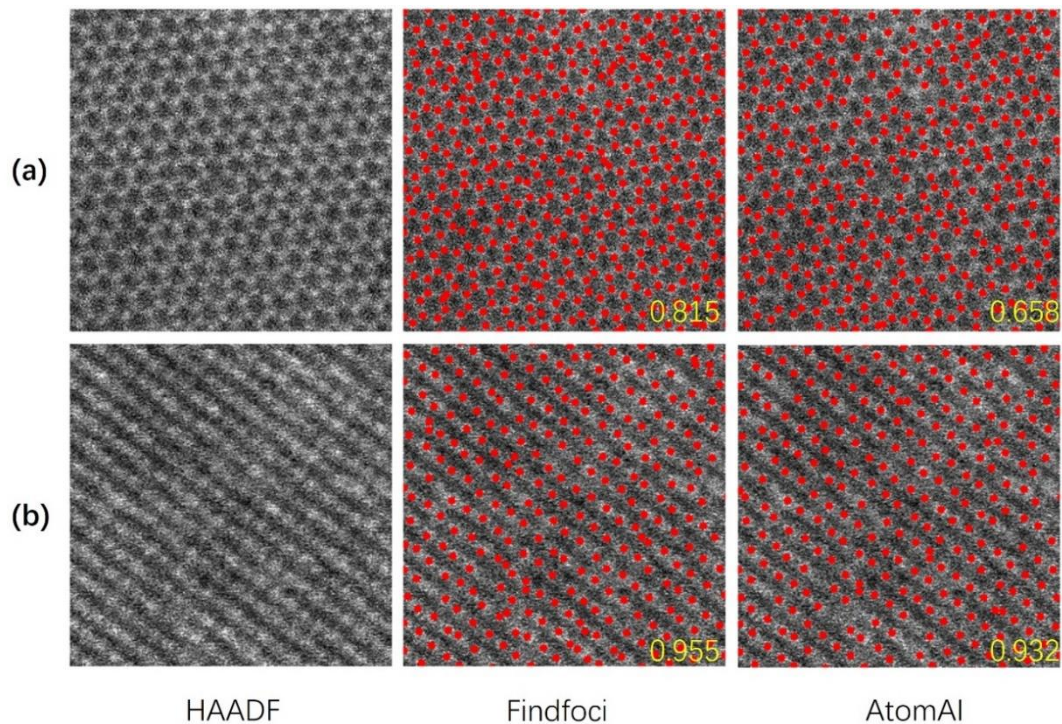

Figure 9. Detection on outliers. The F1-Score is shown in yellow. **(a)** HAADF image of bilayer MoS_2. **(b)** HAADF image of BiSeO. Atomic scale: **(a)** 3.18nm × 3.18nm. **(b)** 3.18nm × 3.18nm.

Furthermore, when faced with unknown structures (not included in the model training), the model also produces relatively poor results, as shown in the **Figure 10**. In this STEM image, there are numerous isolated single atoms/atom clusters that do not form a tightly packed lattice structure. This is unfamiliar to the AtomID model, and despite achieving an accuracy of 77% in the identification results, it generates a significant number of false positive errors. These errors mainly stem from the influence of image noise, leading to misidentification by the model. Therefore, currently, the AtomID model exhibits weak robustness to different structural materials and lacks extrapolation capabilities.

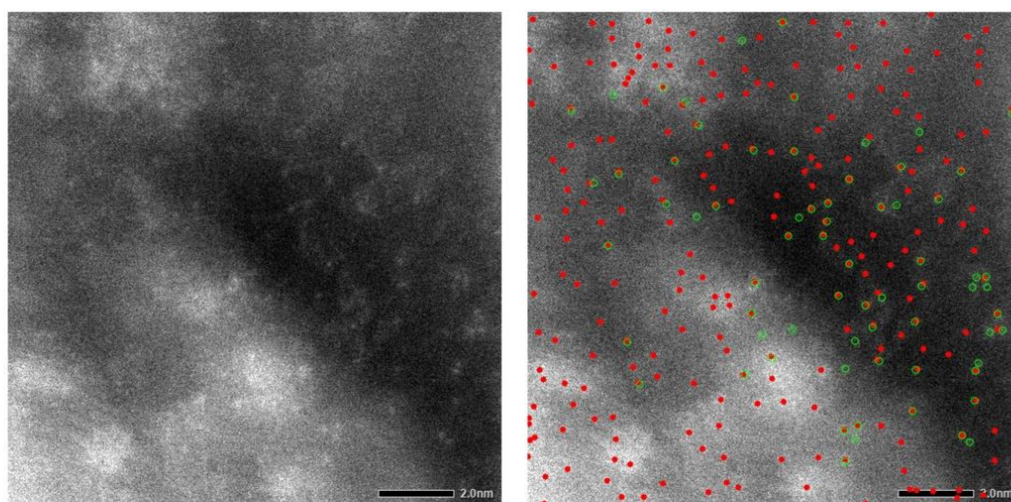

Figure 10. Model performance for aperiodic atomic detection tasks. The green hollow circle is artificially marked, and the red solid circle is the result of the model. Of the 70 points manually marked, 54 atoms were successfully identified. The precision is about 77%. However, there are a lot of false positive errors in the model.

3.5. Strain Mapping

To further evaluate the performance, we used the Findfoci, Atomap, AtomSegNet and our AtomID-net model to find the position of the atomic columns and calculate the strain map of the image. We also used geometric phase analysis (GPA) method [85] to measure two-dimensional strain mapping and use it as a reference image. Here we consider the normal strain mainly, therefore we calculated the hydrostatic normal strain [86] after extracting the atomic positions by the atom identification models. For the test image, we chose an area without any atomic defects (confirmed by high quality STEM image at the same sample location) of MoS_2 and we took the low-quality STEM images for testing our model. By analyzing the high-quality images, there should be very low stain variation in this sample area. By comparing the strain mapping results calculated by the atomic positions extracted from different models and by GPA strain analysis, we found that the result of our AtomID-net (-0.00141 ± 0.0265) is most close to the result of GPA (-0.0178 ± 0.0132), exhibiting the good performance of our model on noisy images **(Figure 11)**.

Figure 11. Strain map calculation and analysis. **(a)** Test image. Atomic scale: 12.7nm

× 12.7nm. **(b)** The normal strain mapping calculated by GPA method. **(c)** Identification result and normal strain mapping calculated by AtomID. **(d)** Identification result and normal strain mapping calculated by Atomap. **(e)** Identification result and normal strain mapping calculated by Findfoci. **(f)** Identification result and normal strain mapping calculated by AtomSegNet. **(g)** The distribution range of the 5 strain mappings above.

3.6. Time Cost

We degenerated the proposed model trained on real data by replacing the NMS module with non-learning detection algorithms: Laplacian of Gaussian (LoG), Difference of Gaussian (DoG) and HOG (Histogram of Oriented Gridients) from scikit-image. The max sigmas are set to the root square of the average spacing with an upper bound 10, to match the spacing parameter in NMS. The thresholds are 0.1, 0.1, 0.004 respectively. Other parameters are set to defaults. Results are shown in **Table 3.**

Table 3. Time consumed and result.

<table>
<tr><th>Metric</th><th>CD</th><th>Jaccard</th><th>F1</th><th>Time</th></tr>
<tr><td>LoG</td><td>2.252±1.036</td><td>0.951±0.039</td><td>0.978±0.023</td><td>727.017±140.831 ms/pic</td></tr>
<tr><td>DoG</td><td>2.320±1.160</td><td>0.963±0.0307</td><td>0.970±0.036</td><td>474.637±139.721 ms/pic</td></tr>
<tr><td>HoG</td><td>4.074±1.336</td><td>0.972± 0.026</td><td>0.939±0.082</td><td>1.192±0.224 s/pic</td></tr>
<tr><td>Ours(cpu)</td><td rowspan="2">2.373±1.176</td><td rowspan="2">0.967±0.025</td><td rowspan="2">0.964±0.039</td><td>231.903±44.965 ms/pic</td></tr>
<tr><td>Ours(gpu)</td><td>52.854 ±2.399 ms/pic</td></tr>
</table>

Our proposed detection module without parameter tuning can exceed the performance of the mentioned strict algorithms that need fine tuning. With the acceleration of GPU, our method has shorter latency and can work in real time.

Chapter 4. Conclusion

Scanning Transmission Electron Microscopy (STEM) is a powerful imaging technique that provides high spatial resolution and detection sensitivity, allowing us to observe and study atomic structures at the atomic scale. This is crucial for improving our understanding of the structure-property relationship in material systems. Therefore, accurate detection of atoms or atomic columns in multiscale and noisy TEM images is essential for material characterization and understanding material properties. However, traditional image processing techniques have limitations when dealing with complex and noisy STEM images. They often struggle with low-quality and highly noisy STEM data.

Deep learning, a branch of machine learning, has shown great potential in the field of STEM image analysis due to its strong feature extraction capabilities. In this paper, we propose an atomic finding model called AtomID-net based on UNet and train and test it using real STEM data. After testing the model on the validation dataset, it achieves an F1 score of 0.964, surpassing traditional image processing techniques and some other deep learning methods. It demonstrates superior performance and flexibility in detecting atomic positions and excels in strain computation, highlighting its accuracy in atomic position identification. Additionally, the model exhibits impressive performance when faced with STEM data at different magnifications. It successfully handles the challenge of images spanning a fourfold increase in scale, ranging from 9 nm to 36 nm, achieving an F1 score close to 0.95, without significant changes in results

due to changes in magnification. This result indicates the robustness of AtomID in handling images at different scales and its adaptability in atomic detection tasks.

Furthermore, we propose a new labeling method for noisy TEM images, where high-quality images are used as references to annotate noisy images of the same sampling positions. This labeling method ensures the acquisition of reliable datasets. Training on real images also enriches the structure and noise variations in the training data, aiding in improving the model's performance in identifying atomic positions in noisy images.

During testing on STEM data with different structures that were not part of the training, we found that the model currently lacks robustness and extrapolation capabilities. In future work, we plan to redesign the model to leverage prior knowledge to enhance performance and improve its robustness.

References

1. Abdollahi, A.; Pradhan, B., Integrating semantic edges and segmentation information for building extraction from aerial images using UNet. *Machine Learning with Applications* **2021,** *6*, 100194.

2. Ahmed, S. F.; Alam, M. S. B.; Hassan, M.; Rozbu, M. R.; Ishtiak, T.; Rafa, N.; Mofijur, M.; Shawkat Ali, A.; Gandomi, A. H., Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review* **2023**, 1-97.

3. Albawi, S.; Mohammed, T. A.; Al-Zawi, S. In *Understanding of a convolutional neural network*, 2017 international conference on engineering and technology (ICET), Ieee: 2017; pp 1-6.

4. Anthony, S. M.; Granick, S., Image analysis with rapid and accurate two-dimensional Gaussian fitting. *Langmuir* **2009,** *25* (14), 8152-8160.

5. Bailer, C.; Habtegebrial, T.; Stricker, D., Fast feature extraction with CNNs with pooling layers. *arXiv preprint arXiv:1805.03096* **2018**.

6. Banhart, J., *Advanced tomographic methods in materials research and engineering*. OUP Oxford: 2008; Vol. 66.

7. Belianinov, A.; Vasudevan, R.; Strelcov, E.; Steed, C.; Yang, S. M.; Tselev, A.; Jesse, S.; Biegalski, M.; Shipman, G.; Symons, C., Big data and deep data in scanning and electron microscopies: deriving functionality from multidimensional data sets. *Advanced Structural and Chemical Imaging* **2015,** *1* (1), 1-25.

8. Binev, P.; Blanco-Silva, F.; Blom, D.; Dahmen, W.; Lamby, P.; Sharpley, R.; Vogt, T., *High-quality image formation by nonlocal means applied to high-angle annular dark-field scanning transmission electron microscopy (HAADF–STEM)*. Springer: 2012.

9. Bourke, P., Cross correlation. *Cross Correlation", Auto Correlation—2D Pattern Identification* **1996**.

10. Caswell, T. A.; Ercius, P.; Tate, M. W.; Ercan, A.; Gruner, S. M.; Muller, D. A., A high-speed area detector for novel imaging techniques in a scanning transmission electron microscope. *Ultramicroscopy* **2009,** *109* (4), 304-311.

11. Cautaerts, N.; Freysoldt, C.; Liebscher, C.; Dehm, G.; Neugebauer, J., Big data and machine learning in electron microscopy. *Interdepartmental research activities–selected highlights*, 222.

12. Chauhan, R.; Ghanshala, K. K.; Joshi, R. In *Convolutional neural network (CNN) for image detection and recognition*, 2018 first international conference on secure cyber

computing and communication (ICSCCC), IEEE: 2018; pp 278-282.

13. Chicco, D.; Jurman, G., The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics* **2020,** *21* (1), 1-13.

14. Comer, M. L.; Delp III, E. J., Morphological operations for color image processing. *Journal of electronic imaging* **1999,** *8* (3), 279-289.

15. Cowley, J., Electron nanodiffraction. *Microscopy research and technique* **1999,** *46* (2), 75-97.

16. Crewe, A.; Wall, J., A scanning microscope with 5 Å resolution. *Journal of molecular biology* **1970,** *48* (3), 375-393.

17. Crewe, A. V.; Wall, J.; Langmore, J., Visibility of single atoms. *science* **1970,** *168* (3937), 1338-1340.

18. Delabie, T.; Schutter, J. D.; Vandenbussche, B., An accurate and efficient gaussian fit centroiding algorithm for star trackers. *The Journal of the Astronautical Sciences* **2014,** *61*, 60-84.

19. Deng, Y.; Bao, F.; Kong, Y.; Ren, Z.; Dai, Q., Deep direct reinforcement learning for financial signal representation and trading. *IEEE transactions on neural networks and learning systems* **2016,** *28* (3), 653-664.

20. Du, X.; Cai, Y.; Wang, S.; Zhang, L. In *Overview of deep learning*, 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), IEEE: 2016; pp 159-164.

21. Duque-Arias, D.; Velasco-Forero, S.; Deschaud, J.-E.; Goulette, F.; Serna, A.; Decencière, E.; Marcotegui, B. In *On power Jaccard losses for semantic segmentation*, VISAPP 2021: 16th International Conference on Computer Vision Theory and Applications, 2021.

22. Egerton, R., Radiation damage to organic and inorganic specimens in the TEM. *Micron* **2019,** *119*, 72-87.

23. Engel, A., Scanning transmission electron microscopy: Biological applications. *Advances in Imaging and Electron Physics* **2009,** *159*, 357-386.

24. Ennos, A., The origin of specimen contamination in the electron microscope. *British Journal of Applied Physics* **1953,** *4* (4), 101.

25. Frank, J., *Electron tomography: methods for three-dimensional visualization of structures in the cell*. Springer: 2008.

26. Gheisari, M.; Wang, G.; Bhuiyan, M. Z. A. In *A survey on deep learning in big data*, 2017 IEEE international conference on computational science and engineering (CSE) and IEEE international conference on embedded and ubiquitous computing

(EUC), IEEE: 2017; pp 173-180.

27. Guo, Y.; Liu, Y.; Oerlemans, A.; Lao, S.; Wu, S.; Lew, M. S., Deep learning for visual understanding: A review. *Neurocomputing* **2016,** *187*, 27-48.

28. Haddad, R. A.; Akansu, A. N., A class of fast Gaussian binomial filters for speech and image processing. *IEEE Transactions on Signal Processing* **1991,** *39* (3), 723-727.

29. He, J.; Li, L.; Xu, J.; Zheng, C., ReLU deep neural networks and linear finite elements. *arXiv preprint arXiv:1807.03973* **2018**.

30. Heider, P.; Pierre-Pierre, A.; Li, R.; Grimm, C. In *Local shape descriptors, a survey and evaluation*, Proceedings of the 4th Eurographics conference on 3D Object Retrieval, 2011; pp 49-56.

31. Herbert, A. D.; Carr, A. M.; Hoffmann, E., FindFoci: a focus detection algorithm with automated parameter training that closely matches human assignments, reduces human inconsistencies and increases speed of analysis. *PloS one* **2014,** *9* (12), e114749.

32. Hinton, G. E., Learning multiple layers of representation. *Trends in cognitive sciences* **2007,** *11* (10), 428-434.

33. Kelleher, J. D., *Deep learning*. MIT press: 2019.

34. Komsa, H.-P.; Kotakoski, J.; Kurasch, S.; Lehtinen, O.; Kaiser, U.; Krasheninnikov, A. V., Two-dimensional transition metal dichalcogenides under electron irradiation: defect production and doping. *Physical review letters* **2012,** *109* (3), 035503.

35. Kong, H.; Akakin, H. C.; Sarma, S. E., A generalized Laplacian of Gaussian filter for blob detection and its applications. *IEEE transactions on cybernetics* **2013,** *43* (6), 1719-1733.

36. Kumar, S.; Kumar, P.; Gupta, M.; Nagawat, A. K., Performance comparison of median and wiener filter in image de-noising. *International Journal of Computer Applications* **2010,** *12* (4), 27-31.

37. Li, H.; Ouyang, W.; Wang, X. In *Multi-bias non-linear activation in deep neural networks*, International conference on machine learning, PMLR: 2016; pp 221-229.

38. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J., A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems* **2021**.

39. Lin, R.; Zhang, R.; Wang, C.; Yang, X.-Q.; Xin, H. L., TEMImageNet training library and AtomSegNet deep-learning models for high-precision atom segmentation, localization, denoising, and deblurring of atomic-resolution images. *Scientific reports* **2021,** *11* (1), 5386.

40. Liu, J., Scanning transmission electron microscopy and its application to the study

of nanoparticles and nanoparticle systems. *Microscopy* **2005,** *54* (3), 251-278.

41. Lu, H.; Li, Y.; Chen, M.; Kim, H.; Serikawa, S., Brain intelligence: go beyond artificial intelligence. *Mobile Networks and Applications* **2018,** *23*, 368-375.

42. Lu, Y.; Chen, Y.; Zhao, D.; Chen, J. In *Graph-FCN for image semantic segmentation*, International symposium on neural networks, Springer: 2019; pp 97-105.

43. Madsen, J.; Liu, P.; Kling, J.; Wagner, J. B.; Hansen, T. W.; Winther, O.; Schiøtz, J., A deep learning approach to identify local structures in atomic-resolution transmission electron microscopy images. *Advanced Theory and Simulations* **2018,** *1* (8), 1800037.

44. Madsen, J.; Susi, T., The abTEM code: transmission electron microscopy from first principles. *Open Research Europe* **2021,** *1* (24), 24.

45. Madsen, J.; Susi, T., ab TEM: A Fast and Flexible Python-based Multislice Simulation Package for Transmission Electron Microscopy. Oxford University Press US: 2023.

46. McNally, K.; Ramirez, D.; Anton, A. M.; Smith, D.; Dick, J. In *Artificial intelligence for space resident objects characterisation with lightcurves*, 8th European Conference on Space Debris, 2021; p 85.

47. Mendes, R. G.; Pang, J.; Bachmatiuk, A.; Ta, H. Q.; Zhao, L.; Gemming, T.; Fu, L.; Liu, Z.; Rümmeli, M. H., Electron-driven in situ transmission electron microscopy of 2D transition metal dichalcogenides and their 2D heterostructures. *ACS nano* **2019,** *13* (2), 978-995.

48. Meyer, J. C.; Eder, F.; Kurasch, S.; Skakalova, V.; Kotakoski, J.; Park, H. J.; Roth, S.; Chuvilin, A.; Eyhusen, S.; Benner, G., Accurate measurement of electron beam induced displacement cross sections for single-layer graphene. *Physical review letters* **2012,** *108* (19), 196102.

49. Muller, D. A., Structure and bonding at the atomic scale by scanning transmission electron microscopy. *Nature materials* **2009,** *8* (4), 263-270.

50. Müller-Caspary, K.; Krause, F. F.; Grieb, T.; Löffler, S.; Schowalter, M.; Béché, A.; Galioit, V.; Marquardt, D.; Zweck, J.; Schattschneider, P., Measurement of atomic electric fields and charge densities from average momentum transfers using scanning transmission electron microscopy. *Ultramicroscopy* **2017,** *178*, 62-80.

51. Neubeck, A.; Van Gool, L. In *Efficient non-maximum suppression*, 18th international conference on pattern recognition (ICPR'06), IEEE: 2006; pp 850-855.

52. Nord, M.; Vullum, P. E.; MacLaren, I.; Tybell, T.; Holmestad, R., Atomap: a new software tool for the automated analysis of atomic resolution images using two-dimensional Gaussian fitting. *Advanced structural and chemical imaging* **2017,** *3*, 1-12.

53. Pennycook, S. J., A scan through the history of STEM. *Scanning Transmission Electron Microscopy: Imaging and Analysis* **2011**, 1-90.

54. Pennycook, S. J., The impact of STEM aberration correction on materials science. *Ultramicroscopy* **2017,** *180*, 22-33.

55. Pennycook, S. J.; Nellist, P. D., *Scanning transmission electron microscopy: imaging and analysis*. Springer Science & Business Media: 2011.

56. Pitas, I.; Venetsanopoulos, A., Nonlinear mean filters in image processing. *IEEE transactions on acoustics, speech, and signal processing* **1986,** *34* (3), 573-584.

57. Postek, M. T., An approach to the reduction of hydrocarbon contamination in the scanning electron microscope. *Scanning: The Journal of Scanning Microscopies* **1996,** *18* (4), 269-274.

58. Pu, Y.; Gan, Z.; Henao, R.; Yuan, X.; Li, C.; Stevens, A.; Carin, L., Variational autoencoder for deep learning of images, labels and captions. *Advances in neural information processing systems* **2016,** *29*.

59. Ravì, D.; Wong, C.; Deligianni, F.; Berthelot, M.; Andreu-Perez, J.; Lo, B.; Yang, G.-Z., Deep learning for health informatics. *IEEE journal of biomedical and health informatics* **2016,** *21* (1), 4-21.

60. Ronneberger, O.; Fischer, P.; Brox, T. In *U-net: Convolutional networks for biomedical image segmentation*, Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer: 2015; pp 234-241.

61. Rouviere, J.-L.; Sarigiannidou, E., Theoretical discussions on the geometrical phase analysis. *Ultramicroscopy* **2005,** *106* (1), 1-17.

62. Shi, W.; Caballero, J.; Theis, L.; Huszar, F.; Aitken, A.; Ledig, C.; Wang, Z., Is the deconvolution layer the same as a convolutional layer? *arXiv preprint arXiv:1609.07009* **2016**.

63. Shibata, N.; Findlay, S. D.; Kohno, Y.; Sawada, H.; Kondo, Y.; Ikuhara, Y., Differential phase-contrast microscopy at atomic resolution. *Nature Physics* **2012,** *8* (8), 611-615.

64. Smith, D. J., The realization of atomic resolution with the electron microscope. *Reports on Progress in Physics* **1997,** *60* (12), 1513.

65. Sun, T.; Gabbouj, M.; Neuvo, Y., Center weighted median filters: some properties and their applications in image processing. *Signal processing* **1994,** *35* (3), 213-229.

66. Thompson, N. C.; Greenewald, K.; Lee, K.; Manso, G. F., The computational limits of deep learning. *arXiv preprint arXiv:2007.05558* **2020**.

67. Tizro, P.; Choi, C.; Khanlou, N., Sample preparation for transmission electron microscopy. *Biobanking: methods and protocols* **2019**, 417-424.

68. Tuceryan, M.; Jain, A. K., Texture analysis. *Handbook of pattern recognition and computer vision* **1993**, 235-276.

69. Ugurlu, O.; Haus, J.; Gunawan, A.; Thomas, M.; Maheshwari, S.; Tsapatsis, M.; Mkhoyan, K., Radiolysis to knock-on damage transition in zeolites under electron beam irradiation. *Physical Review B* **2011,** *83* (11), 113408.

70. Varela, M.; Lupini, A. R.; Benthem, K. v.; Borisevich, A. Y.; Chisholm, M. F.; Shibata, N.; Abe, E.; Pennycook, S. J., Materials characterization in the aberration-corrected scanning transmission electron microscope. *Annu. Rev. Mater. Res.* **2005,** *35*, 539-569.

71. Vaseghi, S. V.; Vaseghi, S. V., Wiener filters. *Advanced Signal Processing and Digital Noise Reduction* **1996**, 140-163.

72. Vijayarani, S.; Vinupriya, M., Performance analysis of canny and sobel edge detection algorithms in image mining. *International Journal of Innovative Research in Computer and Communication Engineering* **2013,** *1* (8), 1760-1767.

73. Vincent, O. R.; Folorunso, O. In *A descriptive algorithm for sobel image edge detection*, Proceedings of informing science & IT education conference (InSITE), 2009; pp 97-107.

74. Von Ardenne, M., Das Elektronen-Rastermikroskop: Theoretische Grundlagen. *Zeitschrift für Physik* **1938,** *109* (9-10), 553-572.

75. Waldrop, M. M., What are the limits of deep learning? *Proceedings of the National Academy of Sciences* **2019,** *116* (4), 1074-1077.

76. Wu, H.; Zhao, X.; Song, D.; Tian, F.; Wang, J.; Loh, K. P.; Pennycook, S. J., Progress and prospects of aberration-corrected STEM for functional materials. *Ultramicroscopy* **2018,** *194*, 182-192.

77. Wu, T.; Pan, L.; Zhang, J.; Wang, T.; Liu, Z.; Lin, D., Density-aware chamfer distance as a comprehensive metric for point cloud completion. *arXiv preprint arXiv:2111.12702* **2021**.

78. Xing, F.; Xie, Y.; Su, H.; Liu, F.; Yang, L., Deep learning in microscopy image analysis: A survey. *IEEE transactions on neural networks and learning systems* **2017,** *29* (10), 4550-4568.

79. Zandvliet, H. J.; van Houselt, A., Scanning tunneling spectroscopy. *Annual review of analytical chemistry* **2009,** *2*, 37-55.

80. Zhang, F.; Li, P.; Hou, Z.-G.; Chen, Y.; Xu, F.; Hu, J.; Li, Q.; Tan, M. In *SEMG feature extraction methods for pattern recognition of upper limbs*, The 2011 international conference on advanced mechatronic systems, IEEE: 2011; pp 222-227.

81. Zhao, F.; Huang, Q.; Gao, W. In *Image matching by normalized cross-correlation*, 2006 IEEE international conference on acoustics speech and signal processing proceedings, IEEE: 2006; pp II-II.

82. Zhao, J.; Deng, Q.; Bachmatiuk, A.; Sandeep, G.; Popov, A.; Eckert, J.; Rümmeli, M. H., Free-standing single-atom-thick iron membranes suspended in graphene pores. *Science* **2014,** *343* (6176), 1228-1232.

83. Zhou, W.; Thomas, J. M., HRTEM surface profile imaging of solids. *Current Opinion in Solid State and Materials Science* **2001,** *5* (1), 75-83.

84. Ziatdinov, M.; Dyck, O.; Maksov, A.; Hudak, B. M.; Lupini, A. R.; Song, J.; Snijders, P. C.; Vasudevan, R. K.; Jesse, S.; Kalinin, S. V., Deep analytics of atomically-resolved images: manifest and latent features. *arXiv preprint arXiv:1801.05133* **2018**.

85. Ziatdinov, M.; Dyck, O.; Maksov, A.; Li, X.; Sang, X.; Xiao, K.; Unocic, R. R.; Vasudevan, R.; Jesse, S.; Kalinin, S. V., Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations. *ACS nano* **2017,** *11* (12), 12742-12752.

86. Zuo, J.-M.; Shah, A. B.; Kim, H.; Meng, Y.; Gao, W.; Rouviére, J.-L., Lattice and strain analysis of atomic resolution Z-contrast images based on template matching. *Ultramicroscopy* **2014,** *136*, 50-60.