Segmentation of Nano-Particles from SEM Images Using Transfer Learning and Modified U-Net

Sowmya Sanan V[1]*, Rimal Isaac R S[2]
Research Scholar, Department of Nanotechnology, Noorul Islam Centre for Higher Education, Kumaracoil, Thuckalay, Tamil Nadu, India[1]
Assistant Professor, Department of Nanotechnology, Noorul Islam Centre for Higher Education, Kumaracoil, Thuckalay, Tamil Nadu, India[2]

***Abstract*—Nanomaterials, owing to their distinctive features, are crucial across numerous scientific domains, especially in materials science and nanotechnology. Precise segmentation of Scanning Electron Microscope (SEM) images is essential for evaluating attributes such as nanoparticle dimensions, morphology, and distribution. Conventional image segmentation techniques frequently prove insufficient for managing the intricate textures of SEM images, resulting in a laborious and imprecise process. In this research, a modified U-Net architecture is presented to tackle this challenge, utilizing a ResNet50 backbone pre-trained on ImageNet. This model utilizes the robust feature extraction abilities of ResNet50 alongside the effective segmentation performance of U-Net, hence improving both accuracy and computational efficiency in TiO_2 nanoparticle segmentation. The suggested model was assessed using performance metrics including accuracy, precision, recall, IoU, and Dice Coefficient. The results indicated a high segmentation accuracy, demonstrated by a Dice score of 0.946 and an IoU of 0.897, with little variability reflected in standard deviations of 0.002071 and 0.003696, respectively, over 200 epochs. The comparison with existing methods demonstrates that the proposed model surpasses previous approaches by attaining enhanced segmentation accuracy. The modified U-Net design serves as an excellent technique for accurate nanoparticle segmentation in SEM images, providing substantial enhancements compared to traditional approaches. This progress indicates the model's potential for wider applications in nanomaterial research and characterization, where precise and efficient segmentation is essential for analysis.**

Keywords*—Nanomaterial; segmentation; ResNet 50; modified UNet; transfer learning; SEM

I. INTRODUCTION

In recent years, nanotechnology has emerged as a revolutionary domain with extensive applications across various industries, including healthcare, energy, electronics, and materials research [1]. Fundamental to several innovations is the capacity to generate and evaluate nanoparticles. Nanoparticles, characterized by at least one dimension ranging from 1 to 100 nanometres, demonstrate distinctive physical and chemical properties that differ from those of their bulk equivalents, attributable to their diminutive size and extensive surface area [2]. Nanoparticles have emerged as a fundamental element of contemporary science and technology, attributable to their distinctive qualities stemming from quantum mechanical phenomena, elevated surface-to-volume ratios, and size-dependent characteristics. These features allow nanoparticles to demonstrate improved optical, mechanical, magnetic, and catalytic characteristics, rendering them essential for various applications.

In medicine, nanoparticles function as drug delivery vehicles, imaging agents for diagnostics, and tools for tissue engineering [3]. Gold nanoparticles are utilized for targeted drug administration and cancer therapy, whereas silver nanoparticles exhibit strong antibacterial traits, rendering them advantageous in medical equipment and wound dressings. Nanoparticles are employed in energy storage to augment the performance of batteries and supercapacitors by improving conductivity and storage capacity. In materials science, nanoparticles are often incorporated into bulk materials to improve strength, flexibility, and thermal characteristics.

Nevertheless, these advances present the issue of precisely describing and assessing nanoparticles, especially regarding their size, shape, distribution, and surface morphology. To conduct a structural characterization of the particles, electron microscopy (EM) is one of the most commonly employed techniques [4]. A particle-interacting electron beam is used in this type of microscope. The objects are reconstructed via the analysis of these interactions and the signals they generate. Transmission electron microscopy (TEM) and SEM are the two primary methods. The primary distinctions between the two technologies are the resolution and the output dimensions. Though TEM images show 2D projections of objects' interior structures, 3D surface reconstruction from SEM images offers important insights into micro/nanoscopic surfaces. The magnification and resolution of TEM surpass those of SEM. Consequently, the SEM is typically employed for morphological characterization, whilst the TEM is utilized for assessing particle size and size distribution, along with other analyses such as phase composition and crystal structure [5].

Analyzing SEM images is a typical procedure used to investigate the outcomes of nanomaterial fabrication processes. By identifying and measuring objects, materials scientists can obtain important morphological information about the material of interest. Image processing procedures for SEM imaging were often executed based on the distinctive qualities of a specific material, including shape, size, brightness, and contrast variations among the observed objects [6]. Nonetheless, the interpretation and analysis of SEM images, particularly for extensive datasets, necessitate effective segmentation techniques capable of precisely delineating nanoparticle

boundaries from intricate backgrounds, and extracting characteristics including dimensions, form, and alignment [7].

The results of segmentation directly affect the precision of subsequent operations, such as nanoparticle counting, size distribution assessment, and morphological investigations. Traditionally, nanotechnologists have depended on manual particle measurement utilizing technologies such as ImageJ, which, while successful for small datasets, becomes labor-intensive and susceptible to discrepancies in large-scale image analysis. Conventional techniques, like template matching, edge detection, and feature extraction-based categorization (utilizing Neural Networks or Support Vector Machines), have proven effective but frequently encounter challenges in generalization as image acquisition methodologies vary. These methods necessitate extensive parameter modifying by experts, making them rigid and time-consuming.

Consequently, conventional techniques frequently fail to yield reliable segmentation outcomes, resulting in erroneous nanoparticle analysis. The drawbacks of current techniques underscore the necessity for a more effective, automated approach to nanoparticle segmentation in SEM images. Recent developments in deep learning have demonstrated significant potential in tackling these challenges. Convolutional Neural Networks (CNNs) have transformed image analysis by facilitating automatic feature extraction and end-to-end learning [8]. CNNs are adept in nanoparticle segmentation tasks, as they can incorporate complex, hierarchical characteristics from raw image data, enabling the collection of subtle details such as nanoparticle borders and textures.

CNN-based models, like U-Net, have shown useful in medical image segmentation; nevertheless, their use for nanoparticle segmentation in SEM images is still inadequately investigated. Moreover, conventional U-Net architectures may inadequately leverage deep learning capabilities for nanoparticle segmentation, as they frequently fail to capture multi-scale features and may experience prolonged training durations when utilized with extensive datasets. This study proposes a novel deep learning (DL)-based segmentation approach to tackle these difficulties. The model is based on the U-Net architecture, known for its effectiveness in biomedical segmentation of images, and integrates a ResNet 50 backbone for improved feature extraction. The incorporation of ResNet50, a robust CNN architecture pre-trained on ImageNet, enables the model to utilize deep residual learning, which has demonstrated efficacy in enhancing the training of extremely deep networks by alleviating the vanishing gradient issue. The main contributions of the proposed research are as follows:

- To introduce a modified U-Net architecture with a ResNet50 backbone, specifically designed for the segmentation of TiO_2 nanoparticles in SEM images.
- By leveraging the pre-trained weights of ResNet50, the model aims to enhance feature extraction capabilities while simultaneously reducing training time and computational resources.
- To highlight the superior performance and applicability of the proposed model nanoparticle analysis compared to existing segmentation techniques.

The subsequent sections of the paper are structured as follows: Section II offers a literature review highlighting existing works and identifying research gaps; Section III elaborates on the proposed model; Section IV discusses the results obtained from the study; A discussion is provided in Section V and finally, a summary of the findings is included in Section VI, which gives a conclusion to the paper.

II. Literature Review

Henrik Eliasson et al. (2024) [9] developed a methodology utilizing two U-Net topologies to independently detect and categorize atomic columns at particle-support interfaces in STEM data. This technique sought to alleviate the problem of noisy data caused by the quick scan speeds required for monitoring nanoparticle movement. The U-Net model was trained on simulated non-physical images and was assessed in comparison to established solutions like AtomSegNet and AtomAI, exhibiting superior performance in both in situ and ex situ time series of diminutive Pt nanoparticles on CeO_2. Experimental time series, captured at 5 frames per second, exhibited dynamic, site-specific displacement of atomic columns. Model training and evaluation were performed on an NVIDIA RTX 4090 GPU, necessitating around 40 minutes for localization and three hours for segmentation. The findings highlighted the method's reliability and precision in examining dynamic nanoparticle behavior.

Nina Gumbiowski et al. (2023) [10] performed an investigation of metallic nanoparticles utilizing a machine learning approach that focused on segmentation, the differentiation of overlapping particles, and individual identification. An approach employing ultimate erosion of convex shapes (UECS) was devised to address particle overlap, enabling the assessment of characteristics such as size, circularity, and Feret diameter within a large particle population. The automated analysis of TEM images successfully extracted shape- and size-related data, including that of nanoscale gold nanoparticles. They employed a DL model, namely a deeplabv3+ network with a ResNet-18 backbone, demonstrating that their method sustained performance over diverse contrast levels, surpassing traditional image processing techniques. This automation markedly decreased analysis duration relative to manual techniques and enabled the extraction of extensive data on particle attributes. Nevertheless, constraints remained in the analysis of significantly overlapping particles, highlighting the difficulties in precisely understanding two-dimensional projections in microscopy.

Jonas Bals and Matthias Epple (2023) [11] employed CNN to independently interpret nanoparticle images acquired from scanning electron microscopy. A framework was built for obtaining secondary electron (SE) and STEM images, enabling quantitative studies of particle size and shape. Utilizing pixel weight loss maps to train CNNs enhanced the segmentation of overlapping particles. Their approaches encompassed the classification of forms, including cubes and spheres, the segregation of particles, and the removal of agglomerates. They attained great accuracy in shape classification utilizing AlexNet and ResNet34, together with efficient segmentation employing UNet++. Nevertheless, the system encountered

difficulties with particles that displayed ambiguous forms or partial obscurity, resulting in a considerable misclassification rate. Moreover, challenges associated with particle overlap and intensity fluctuations in SE images further impeded the efficacy of their approach.

Matthew Helmi Leth Larsen et al. (2023) [12] examined the most effective frame dose for object detection and segmentation in low electron dose TEM imaging. The MSD-net architecture exhibited superior performance compared to the regular U-net, particularly at frame dosages lower than those utilized during training. The MSD-net successfully segmented Au nanoparticles, attaining visibility at dosages as low as 20–30 e−/Å² and full segmentation at 200 e−/Å². During training with simulated images, the study highlighted how crucial it is to model the modulation transfer function (MTF), hence improving the network's capacity to identify nanoparticles in low signal-to-noise ratio environments. Through benchmarking the U-net and MSD-net across different frame dosages, the authors demonstrated that the MSD-net can generalize beyond its training data, establishing it as the best option for analyzing noisy images.

Wenkai Fu et al. (2022) [13] developed a machine learning model for predicting temporal sequences of in-situ TEM video frames with a hybrid long-short-term memory (LSTM) algorithm and a features de-entanglement technique. They used deep learning algorithms to predict future video frames from prior ones, offering insights into size-dependent structural alterations in Au nanoparticles under dynamic response settings. The models exhibited notable performance, with a structural similarity value of roughly 0.7, while being trained on limited datasets. The researchers accurately forecasted the shifts of Au nanoparticles from rigid to dynamic forms, underscoring the model's relevance in catalytic science. Although the structural similarity scores were inferior to those from more extensive benchmark datasets, the PhyDNet model proficiently delineated the structural evolution of Au nanoparticles.

Zhongyuan Ji and Yuchen Wang (2022) [14] utilized TEM to examine and evaluate the morphological characteristics of nanoparticles. A multirandom forest algorithm for image segmentation was devised, which markedly surpassed conventional techniques like maximum entropy threshold segmentation and watershed segmentation. Utilizing FCM clustering and the algorithm's ability to manage diverse gray levels in TEM images, they attained segmentation accuracy of up to 95%. Notwithstanding these gains, the methodology exhibited certain limitations, such as sensitivity to fluctuations in sample characteristics and image quality, along with computing difficulties in handling extensive datasets.

Annick De Backer et al. (2022) [15] presented a Bayesian genetic approach to reconstruct atomic models of monotype crystalline nanoparticles from single projections utilizing Z-contrast imaging. They employed atom-counting data from annular dark field STEM images as input for preliminary 3D models. The approach reduced structural energy while integrating past knowledge regarding atom-counting accuracy and neighbor-mass relationships. The results indicated enhanced reliability in reconstructing beam-sensitive nanoparticles, especially those approximately 3 nm in size, at low electron doses. Their comprehensive simulation analysis objectively assessed the reconstructions, demonstrating a substantial improvement in the precise identification of surface atoms. The study demonstrated that the incorporation of finite atom-counting precision and neighbor-mass relationships significantly enhanced the quality of 3D atomic models, indicating improved predictions of catalytic capabilities for future applications. The algorithm was subsequently utilized to evaluate a time series of experimental photographs of a platinum nanoparticle, demonstrating its efficacy in real-time structural measurement.

Hari Mohan Singh et al. (2022) [16] developed a machine learning regression model, Gradient Boost Regression (GBR), to predict the particle size of aluminum nitride (Al_2N_3), silicon nitride (Si_3N_4), and titanium nitride (TiN) nanoparticles dispersed in ethylene glycol (EG) solution. They found critical factors affecting density, including nanoparticle size, molecular weights, volume concentration, and temperature. The GBR model demonstrated significant accuracy in forecasting nanofluid density for a training dataset, nearly matching experimental values obtained from a DMA 500 density meter across different temperatures and concentrations. The research utilized Gradient Search Optimization (GSO) for hyperparameter optimization to improve model performance. The outcomes show a robust correlation between the GBR predictions and experimental data, highlighting the model's efficacy.

Alexey G. Okunev et al. (2020) [17] investigated the automated identification of metal nanoparticles on highly oriented pyrolytic graphite utilizing deep learning methodologies, particularly the Cascade Mask-RCNN neural network. Their model was trained on a dataset including 23 scanning tunnelling microscopy (STM) images, which included 5,157 labeled nanoparticles, and was subsequently validated on a distinct set of images. The trained network achieved a precision of 0.93 and a recall of 0.78, illustrating its proficiency in identifying nanoparticles with an accuracy range of 0.87–0.99 for mean particle size assessments. The study emphasized the limitations of traditional image processing techniques, which required high-quality images and significant parameter adjustment. Researchers have created the open-access web service "ParticlesNN," enabling researchers to analyze noisy STM images without prior improvement, thereby markedly enhancing nanoparticle identification and quantification across diverse imaging scenarios.

Horwath et al. (2019) [18] examined sophisticated deep learning techniques for the segmentation of nanoparticles in EM images. They discovered that although the speed and quality of image segmentation had enhanced, the application of deep learning methods to precisely capture physical attributes continued to pose difficulties. The model's generalization was limited by the necessity for pixel-level annotations and the class imbalance in the training datasets, necessitating meticulous preparation. The effectiveness of segmentation on high-resolution images was further compromised by noise and light fluctuations, requiring more focus during training. Their experiments with various CNN configurations highlighted the importance of batch normalization and kernel size in enhancing

model accuracy and stability. Nevertheless, problems persisted, including sensitivity to fluctuations in image resolution and a tendency to overfit the training data.

Yi-Chi Wang et al. (2019) [19] examined the issues associated with employing STEM tomographic imaging to delineate 3D elemental segregation in nanoparticles, mainly arising from electron dose constraints in conventional techniques. They used a method known as spectroscopic single particle reconstruction (SPR), derived from structural biology, to assess PtNi nanocatalysts. They effectively identified nanoparticles with a diameter of 20 ± 2 nm and a platinum concentration of 56 ± 6 atom% by integrating both STEM-EDS and STEM-HAADF images. The significant diversity in nanoparticle size, shape, and composition required rigorous selection criteria for accurate characterization in the SPR technique.

Although methods for segmenting and detecting nanoparticles in electron microscopy have advanced, there are still a number of significant drawbacks. Numerous current methodologies, including deep learning techniques, encounter difficulties in precisely segmenting overlapping nanoparticles, particularly under fluctuating imaging settings like low signal-to-noise ratios and elevated electron doses. Conventional segmentation techniques often rely on extensive pixel-level annotations and encounter class imbalance, leading to overfitting and insufficient generalization across diverse datasets. Furthermore, traditional algorithms can inadequately consider the complexities of nanoparticle morphology, such as irregular shapes and agglomeration effects, limiting classification accuracy. Dependence on certain imaging modalities sometimes limits these models' ability to include diverse types of electron microscopy data. To address these limitations, we present a novel methodology designed to enhance segmentation accuracy and robustness across various nanoparticle types and imaging conditions, thus facilitating more precise analysis in materials science and nanotechnology applications.

III. Materials and Methods

The proposed modified U-Net architecture, incorporating a ResNet50 backbone, was chosen due to its superior feature extraction capabilities, which are critical for accurately segmenting TiO_2 nanoparticles in SEM images. The deep residual learning framework of ResNet50 enables the model to effectively capture intricate textural details and morphological variations, thereby addressing the limitations of conventional segmentation approaches that struggle with complex nanoparticle structures. Furthermore, the integration of ResNet50 with U-Net enhances segmentation accuracy while maintaining computational efficiency, leveraging pre-trained ImageNet weights to expedite training and improve generalization across diverse SEM datasets. In contrast, traditional U-Net models, while effective for biomedical image segmentation, lack the depth required for precise nanoparticle boundary detection, often resulting in suboptimal segmentation performance. Moreover, standard CNN-based approaches necessitate extensive manual feature engineering and fail to adequately handle the high variability in SEM image textures. High-complexity architectures, such as DeepLabV3+, although capable of achieving high accuracy, impose significant computational overhead, making them less practical for real-time applications. The proposed modified U-Net architecture effectively mitigates these challenges, providing a robust, scalable, and precise segmentation framework that significantly outperforms existing methodologies in the domain of nanomaterial characterization.

The approach seeks to improve the segmentation of nanoparticles in high-resolution SEM images by incorporating transfer learning and a modified U-Net (mUNet) architecture. A pre-trained ResNet50 model is employed, originally constructed on the ImageNet dataset, to extract significant features from SEM images depicting various nanoparticle shapes, distributions, and sizes. Each image in the dataset is paired with a ground truth mask that delineates the nanoparticles, enabling supervised learning. Pre-processing techniques, such as image scaling and normalization, enable the standardization of inputs for the model. The ResNet50 backbone serves as a feature extractor, acquiring multi-scale, intricate representations of the SEM images, which the mUNet model then refines for accurate segmentation. The efficiency of the model is evaluated following training by accuracy, precision, recall, Intersection over Union (IoU), and Dice Coefficient to confirm its ability to separate nanoparticles under diverse imaging settings. By overcoming the issues related to SEM images, this method aims to increase the precision and efficacy of nanoparticle segmentation. Fig. 1 displays the suggested framework's block diagram.

A. Dataset

The dataset for the proposed study is obtained from the publicly accessible repository on GitHub (https://github.com/BAMresearch/automatic-sem-image-segmentation). The database includes EM micrographs of TiO_2 particles and their related segmentation masks that define the boundaries of these particles. The collection additionally encompasses classifications of the particles according to their visibility and occlusion. The set is organized into subfolders that include raw SEM and TSEM images, along with manually annotated segmentation and classification masks. This extensive dataset establishes a robust basis for training and assessing the mUNet model, ensuring the precision and dependability of the segmentation procedure through the utilization of high-quality, real-world SEM data. Fig. 2 shows the first and last sample SEM images from the dataset along their corresponding manually annotated segmentation mask.

B. Data pre-processing

Pre-processing techniques such as scaling and normalization are employed to normalize the SEM images for uniform input into the neural network. Resizing standardizes images to a consistent dimension, whereas normalization calibrates pixel values to meet the network's specifications, hence improving the model's data processing efficiency. Standardizing the inputs enhances the model's robustness and its capacity to manage changes in nanoparticle morphology within SEM images. The dataset was divided into two categories, with 85% allocated for model training and the remaining 15% allotted for validation. This curated dataset is

essential for training and assessing the mUNet model, enabling precise and resilient segmentation performance.

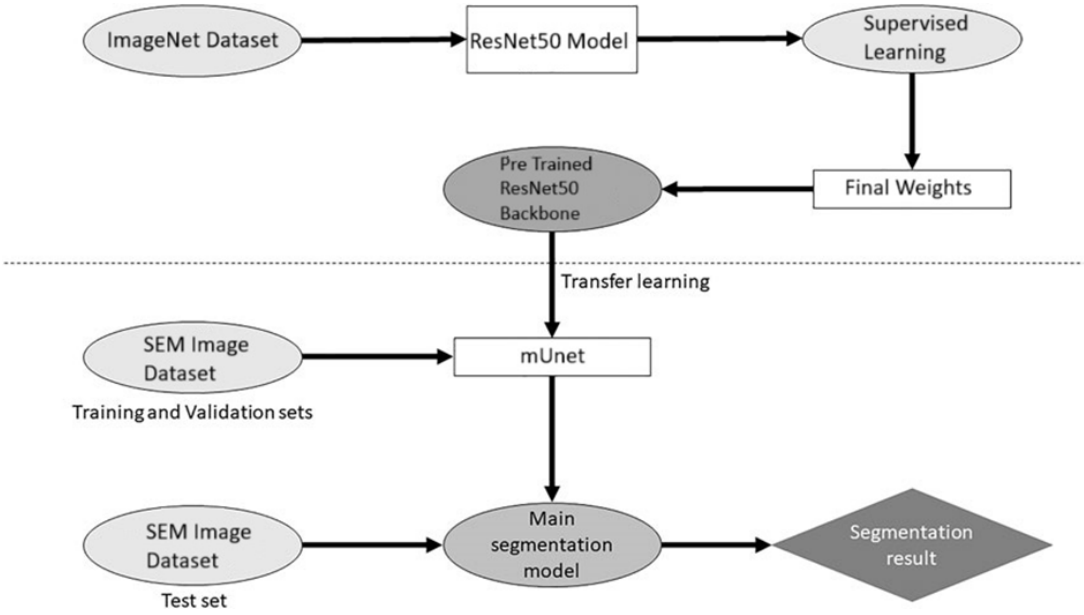


Fig. 1. Block diagram of the proposed model.

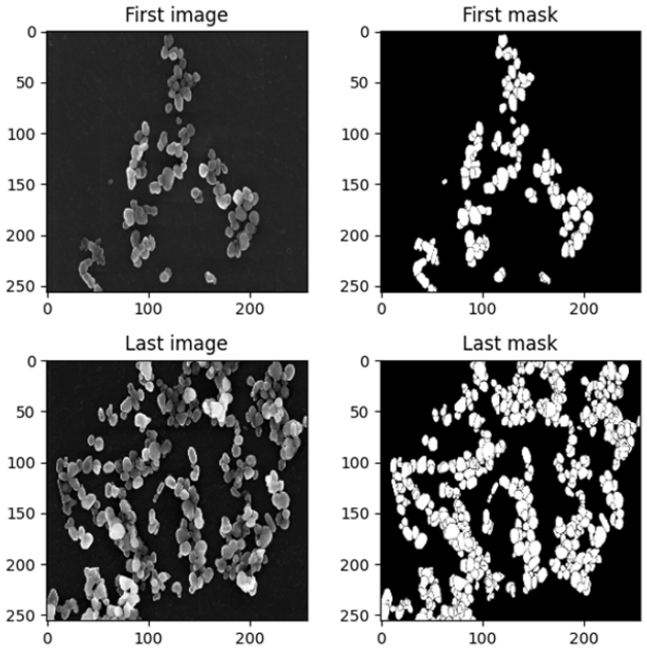


Fig. 2. Sample images from the dataset.

C. Model Development

Segmentation is an essential procedure in image analysis that entails dividing the images into separate sections, facilitating the identification and localization of particular objects within the image. This research utilizes deep learning techniques to improve segmentation accuracy, specifically for the identification of nanoparticles in SEM images. The U-Net design, acclaimed for its efficacy in biomedical image segmentation, underpins the suggested methodology. A mUNet employs a pre-trained ResNet50 encoder to enhance feature extraction and gather multi-scale information. This combination enables the accurate delineation of nanoparticle borders, enhancing both the precision and efficiency of the segmentation process.

1) U-Net: The U-Net is a widely employed CNN model, particularly adept at segmentation tasks, such as nanoparticle segmentation in high-resolution SEM images. UNet is engineered for pixel-level classification, facilitating the differentiation of regions of interest (ROI) from the background in images, which is particularly beneficial in

domains like medical imaging, biological analysis, and materials science [20]. The architecture is characterized by its U-shaped structure, with two main pathways: a contracting pathway and an expansive pathway, as shown in Fig. 3. The contracting path derives feature maps from the input image using many convolutional layers, whereas the expansive path reconstructs these features into a segmented output by integrating low-level and high-level features.

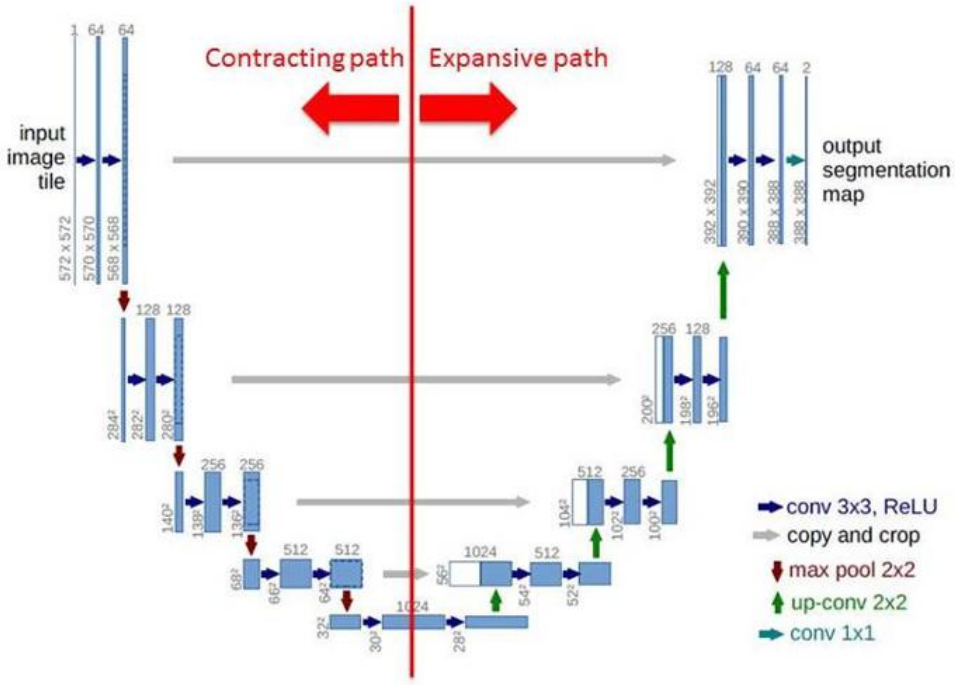


Fig. 3. Basic UNet architecture.

The contracting path, referred to as the encoder, adheres to the conventional architecture of a CNN. The process entails consecutive applications of two 3×3 convolutional layers, succeeded by a ReLU activation function and a 2×2 max-pooling operation. The max-pooling procedure aims to down sample the image, decreasing its spatial dimensions while simultaneously doubling the number of feature channels at each iteration. The convolutional procedure is mathematically expressed as shown in Eq. (1).

$$f(x) = \sigma(W * x + b) \quad (1)$$

where, W denotes the convolutional weights, x signifies the input image, b indicates the bias, and σ represents the activation function.

The expansive path, known as the decoder, mimics the contracting path in reverse, with the objective of reconstructing the image's spatial dimensions while preserving localization accuracy. At each phase of the expanding pathway, the feature maps are subjected to up sampling, generally via a 2×2 transpose convolution process, to revert the image to its original resolution. The transpose convolution is mathematically represented in Eq. (2).

$$\hat{x} = W^T * f(x) \quad (2)$$

Where, W^T denotes the transposed weights utilized in the up-sampling process, while x signifies the upsampled feature map. This step is succeeded by concatenation with the relevant feature maps from the contracting path, enabling the network to merge low-resolution, high-context data with high-resolution, low-context features. The concatenation technique preserves small information from prior layers, which is essential for precise nanoparticle segmentation.

Skip connections are crucial to the performance of the U-Net [21]. These connections directly associate feature maps from the contracting path with the expanded path, as shown in Fig. 4, enabling the model to preserve intricate spatial information that could be lost during downsampling. By integrating low-level, high-resolution characteristics with up sampled high-level features, the network can generate precise and detailed segmentations.

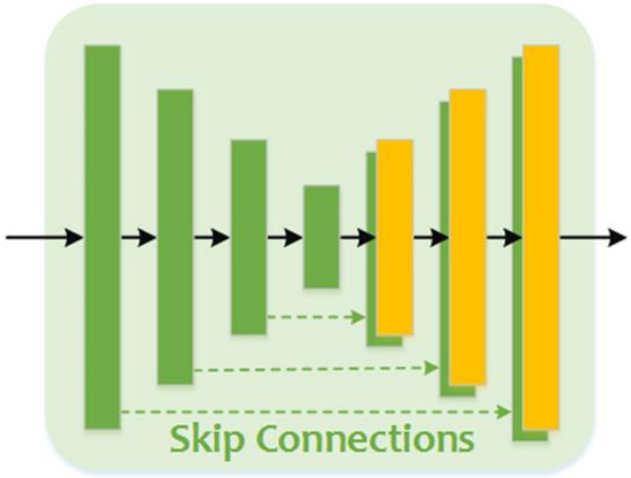


Fig. 4. Skip connections in the UNet architecture.

In the final layer, a 1×1 convolution is utilized to transform the 64-channel feature map into the requisite number of output classes (e.g., foreground and background in binary segmentation), mathematically represented as in Eq. (3). This generates the definitive segmentation mask, categorizing each pixel based on its respective region.

$$y = W_{final} * f(x) + b_{final} \quad (3)$$

Where W_{final} represents the final layer's weights, and b_{final} is the bias term.

A distinctive feature of the U-Net architecture is its capacity to manage huge images via an overlap-tile method. Due to GPU memory limitations on image size, U-Net divides huge images into smaller tiles, processes each tile independently, and then merges them to generate a comprehensive segmentation map. To mitigate the loss of context at the image peripheries, the input image is mirrored throughout the tiling procedure. This technique guarantees precise segmentation of edge pixels, even when analyzing extensive SEM images of nanoparticles.

2) Proposed modified UNet architecture: The proposed research enhances the traditional U-Net design by integrating a modified U-Net structure that employs a ResNet50 encoder. This modification seeks to utilize the powerful feature extraction skills of ResNet50 to enhance nanoparticle segmentation performance from SEM images. The modification concentrates on optimizing feature extraction, refining multi-scale representations, and improving the overall accuracy of segmentation results.

ResNet50 is a CNN that uses residual learning to improve the training of deep networks, facilitating the effective acquisition of intricate hierarchical features vital for precise segmentation tasks. ResNet50 comprises 50 layers organized into many blocks, as shown in Fig. 5, using skip connections, enabling the model to learn residual functions [22]. This architecture mitigates the degradation issue frequently observed in deep networks, wherein greater depth results in reduced performance. ResNet50 mitigates vanishing gradient problems by establishing direct paths for gradients during back propagation, hence enhancing the training of deeper networks. Each residual block in ResNet50 comprises two or three convolutional layers, batch normalization, and ReLU activation algorithms, ending in a shortcut connection that bypasses one or more layers.

In a standard residual block, the input X undergoes a sequence of convolutions, after which the original input is reintegrated into the output, resulting in Eq. (4).

$$Y = F(X) + X \quad (4)$$

Where F(X) denotes the function acquired by the convolutional layers, and Y signifies the output of the block. This architecture enhances the acquisition of identity mappings, hence aiding the training of more profound networks.

The ResNet50 encoder comprises numerous convolutional layers arranged in blocks. Each block generally consists of three convolutional layers: the initial layer applies a 1x1 convolution to diminish dimensionality, the subsequent layer employs a 3x3 convolution for feature extraction, and the last layer utilizes another 1x1 convolution to reinstate the original dimensionality. This setup improves the model's ability to represent spatial hierarchies while preserving computational efficiency. The implementation of batch normalization subsequent to each convolutional layer enhances learning stability by normalizing the inputs to each layer.

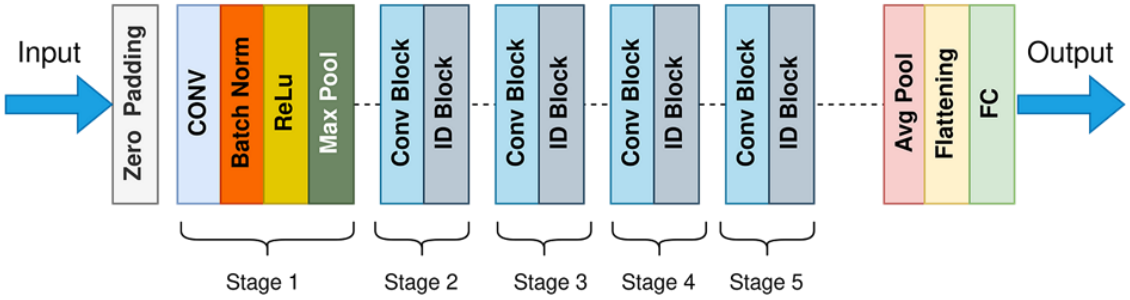


Fig. 5. ResNet 50 architecture.

The downsampling is achieved by strided convolutions and max pooling layers. Strided convolutions diminish the spatial dimensions of feature maps while augmenting depth, efficiently capturing multi-scale characteristics at diverse levels of abstraction. The downsampling process is essential for the encoder, enabling the model to concentrate on more intricate, abstract representations of the input data as it advances through the layers. As the input SEM images traverse the ResNet50 encoder, feature maps are produced at various depths. These feature maps encompass a comprehensive array of attributes that capture both low-level features (like textures and edges) and high-level semantic data (such as patterns and shapes). The hierarchical structure of feature extraction allows the model to acquire intricate representations essential for precisely segmenting nanoparticles in the images. The formula for feature extraction at any specified layer is represented as shown in Eq. (5).

$$F_i = \sigma(W_i . X + b_i) \quad (5)$$

Where F_i denotes the feature map at layer i, W_i and b_i represents the weights and biases of the convolutional layer, whereas X signifies the input feature map from the preceding layer.

The deepest convolutional layers process the smallest and most abstract feature mappings near the architecture's

bottleneck. This component encapsulates the most complicated illustrations of the input data, enabling the network to proficiently discern complex patterns related to nanoparticles. The architectural design enables the model to manage high-level properties while preserving essential spatial information required for precise segmentation.

During the decoder step, the design utilizes transpose convolutions (deconvolutions) to up sample the feature maps to their original input dimensions. The mUNet employs a concatenation method that integrates the up sampled feature maps with the matching encoder feature maps at different levels, rather than simply copying the feature maps. This procedure ensures the preservation and integration of intricate details and spatial data from preceding layers with enhanced semantic attributes. The concatenation is represented mathematically as in Eq. (6).

$$F_{concat} = F_{upsampled} \oplus F_{encoder} \quad (6)$$

Where F_{concat} denotes the concatenated feature map, $F_{upsampled}$ refers to the feature map subsequent to upsampling, and $F_{encoder}$ signifies the feature map derived from the encoder.

The output layer employs a 1×1 convolution to diminish the feature map's channel number to one, appropriate for binary segmentation applications. A sigmoid activation function is utilized on the output to produce a segmentation mask, resulting in values ranging from 0 to 1, which represent the probability of each pixel being part of the target class (nanoparticles). The final segmentation mask is represented as shown in Eq. (7).

$$M = \sigma(W_{output} . F_{concat} + b_{output}) \quad (7)$$

Where M represents the output mask, and W_{output} and b_{output} denote the weights and bias for the output layer.

Post-processing techniques, including thresholding and morphological processes, are utilized to enhance the raw segmentation outcome. The proposed research achieves enhanced segmentation performance due to the creative utilization of convolutional layers, downsampling techniques, and residual connections, which facilitate an in-depth knowledge of the input data.

D. Hardware and Software Setup

The study utilized a high-performance computational configuration comprising an Intel Core i7 CPU, 32GB of RAM, and an NVIDIA GeForce GTX 1080Ti GPU, facilitating the effective management of demanding computational workloads. The framework was executed with the Keras library, a high-level neural network API based on TensorFlow, recognized for its user-friendly interface and robust functionalities. The training procedure was conducted on Google Colab, a cloud-based Python notebook platform that offers easy accessibility to substantial computational resources, hence facilitating model training.

An essential element of this research was the selection of hyperparameters, which profoundly influence model performance during training. Unlike model parameters that are derived from the data hyperparameters are predetermined by the user and are crucial in shaping the configuration of the training process to optimize the performance of the nanoparticle segmentation model. The precise hyperparameter selections and model configuration are detailed in Table I.

TABLE I. HYPERPARAMETER SPECIFICATIONS

Hyper parameters	Values
Epochs	200
Learning Rate	0.0001
Optimizer	ADAM
Batch size	4
Loss function	Dice loss

IV. EXPERIMENTAL RESULTS

Initially, several factors are defined to quantify essential performance parameters, as represented in following Eq. (8), Eq. (9), Eq. (10), Eq. (11), Eq. (12), and Eq. (13). These metrics, based in the principles of False Positive (FP), True Negative (TN), False Negative (FN), and True Positive (TP), are crucial for evaluating the efficacy of the model.

The calculation of accuracy involves dividing the total number of predictions by the number of right predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

The exactness of a prediction is measured by its precision, or the number of true positives. Instead, recall quantifies completeness, or the number of real positives that were anticipated as positives.

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

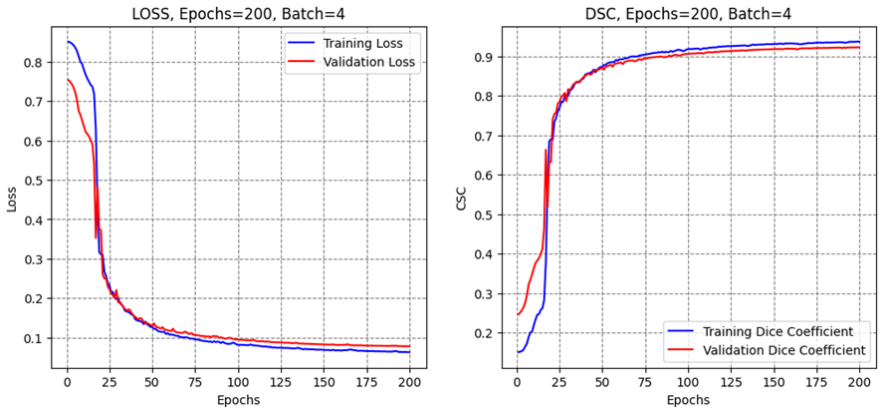
$$F1 - Score = 2 * \left(\frac{Precision * Recall}{Precision + Recall}\right) \quad (11)$$

$$Dice\ Score = \frac{2 \times A \cap B}{|A| + |B|} \quad (12)$$

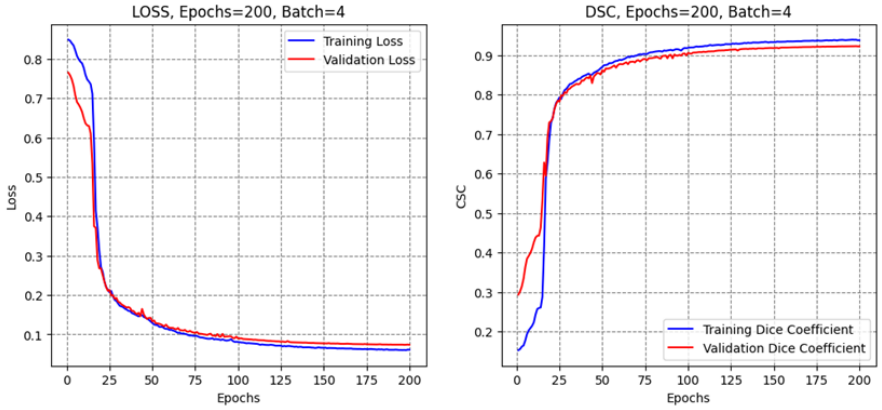
$$IoU = \frac{|A \cap B|}{A \cup B} \quad (13)$$

Where, A and B denote the set of predicted and actual positive instances.

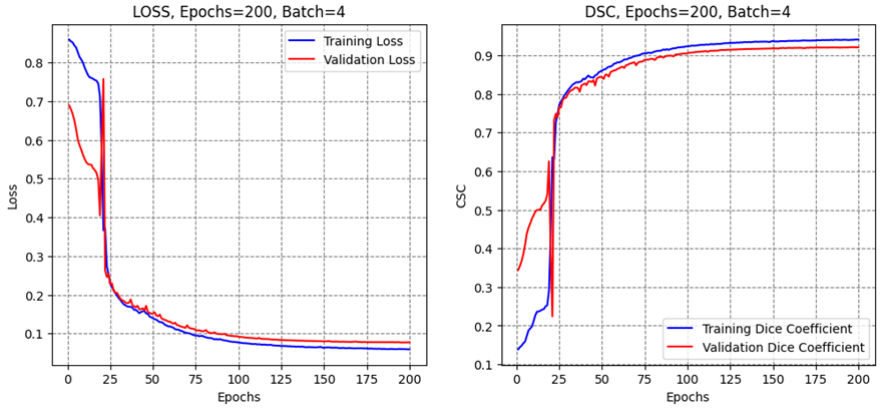
The dataset is partitioned into five folds using KFold, which randomly assigns data to training and validation sets for each iteration. For each fold, the model is established and constructed with five iterations of 200 epochs each. The training procedure includes a 15% validation split, with EarlyStopping callbacks to avert overfitting by ceasing training when the validation loss does not increase over a predetermined number of epochs. Following each training session, the model's learning progression is depicted via the training and validation loss curves, as shown in Fig. 6. Metrics from each iteration are recorded for subsequent analysis, facilitating a thorough assessment of the model's performance across several folds.



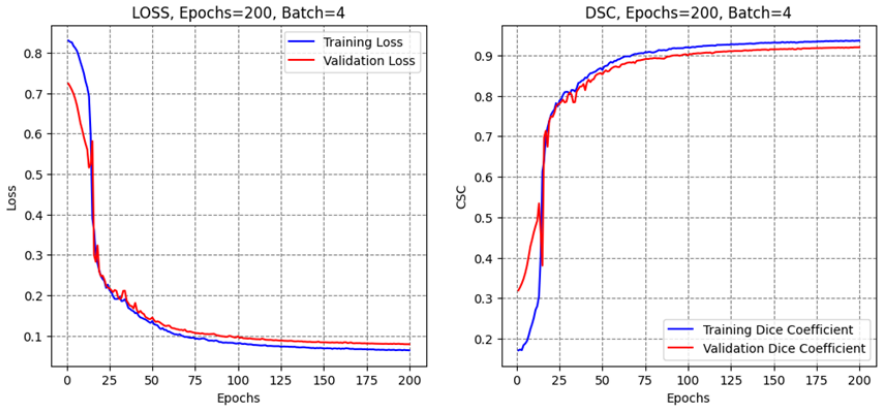
(a) Run 1



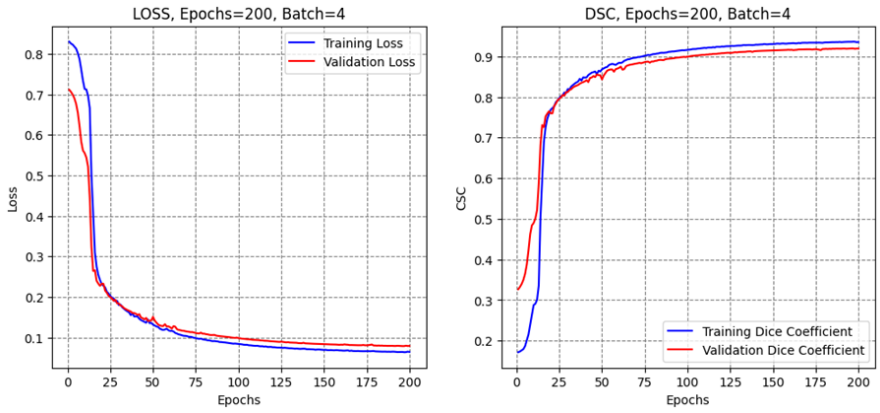
(b) Run 2



(c) Run 3



(d) Run 4



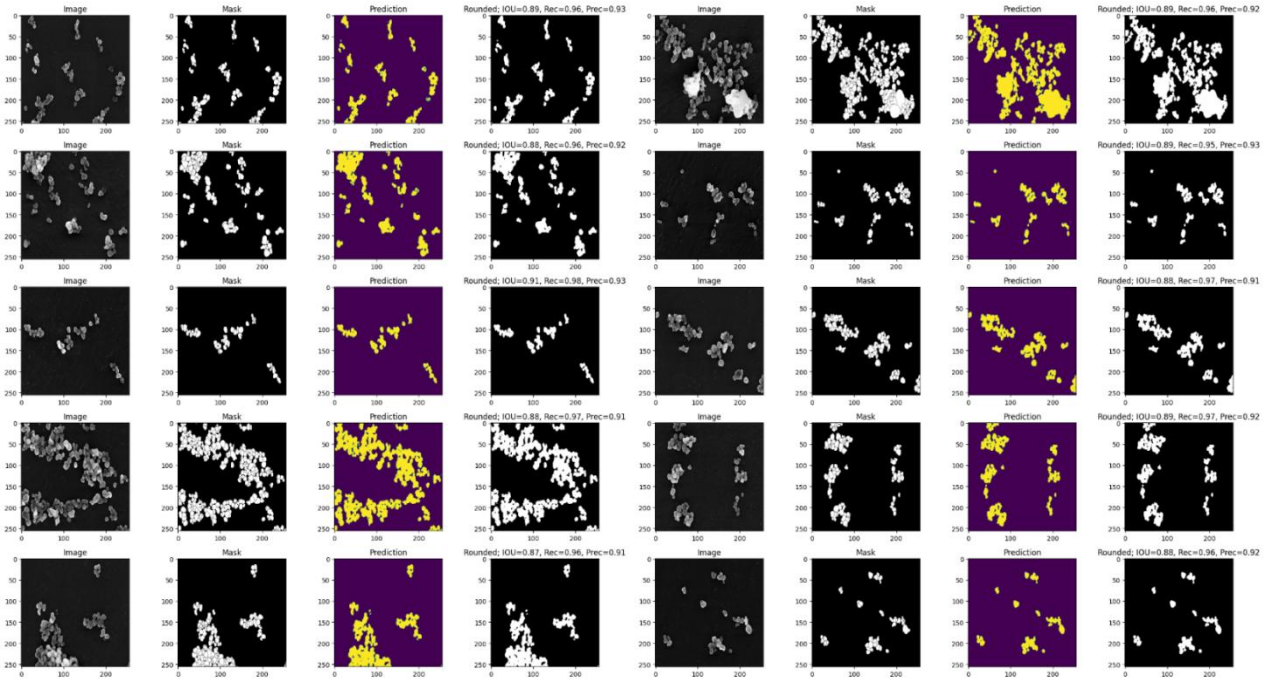
(e) Run 5

Fig. 6. Training and validation dice coefficient of the proposed model.

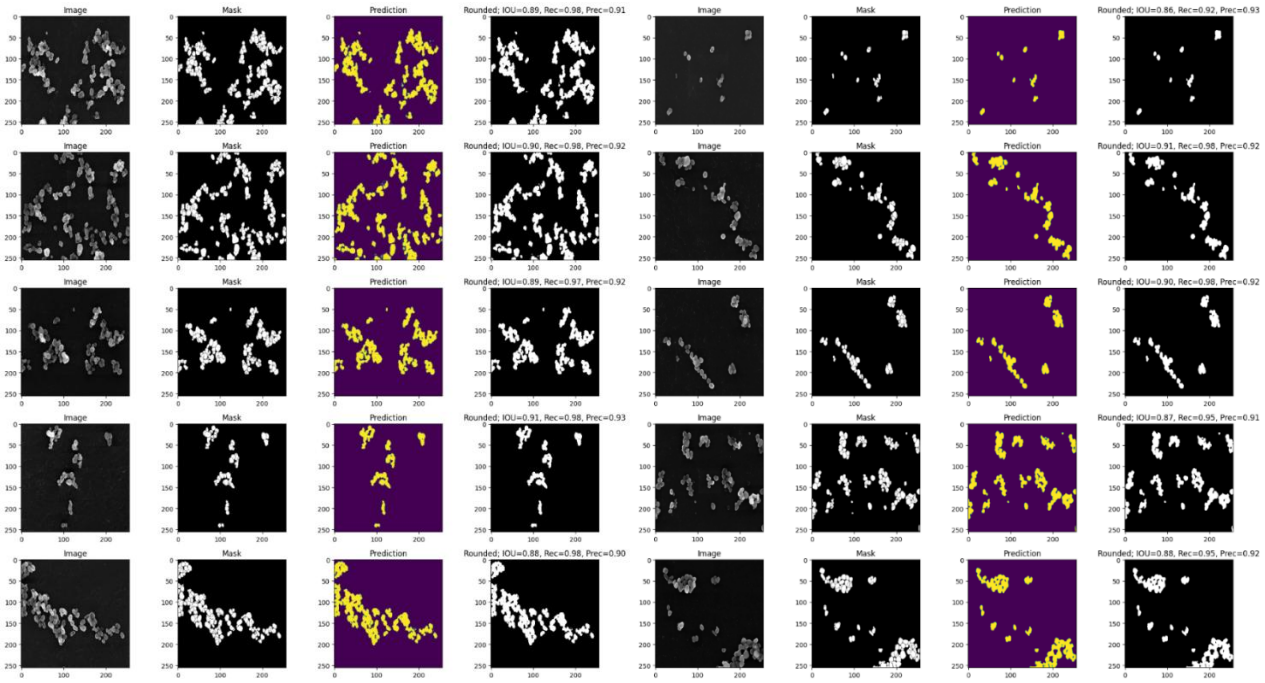
The incorporation of visualizations that compare predicted masks with original images, as illustrated in Fig. 7, improves interpretability and facilitates the evaluation of segmentation quality.

The model's precision metric exhibits a gradual enhancement across the runs, as shown in Fig. 8 (a) and Table II, commencing at 0.907 in the initial run and culminating at 0.926 in the final run. This pattern indicates the model's enhanced capacity to accurately identify true positive nanoparticle boundaries. Notwithstanding a slight decline to 0.900 in the third run, overall precision enhanced, especially in the fourth run, where it attained 0.916. The recall scores consistently stayed elevated during the runs, as shown in Fig. 8 (b), starting at 0.974, decreasing slightly to 0.969 in the second run, and achieving a maximum of 0.979 in the third run. The most recent test registered a decrease to 0.966; however, the recall metrics continually demonstrated the model's capability in detecting nanoparticle segments.

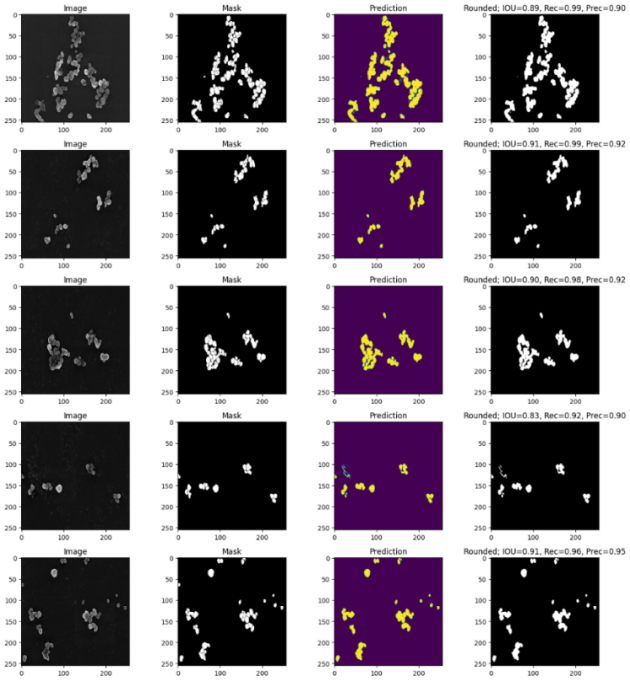
The IoU score demonstrated robust performance, as shown in Fig. 8 (c), commencing at 0.885, with minor fluctuations during the second and third runs, and ultimately increasing to 0.897 in the last run. This enhancement indicates improved precision in delineating nanoparticle boundaries. The DSC commenced at 0.938, as shown in Fig. 8 (d), and reached a peak of 0.946 in the third run, indicating the highest level of segmentation accuracy attained. A decrease in DSC was noted in consecutive trials, culminating in a final value of 0.940. The mUNet model demonstrates robust segmentation capabilities, with significant performance enhancements across multiple parameters, confirming its efficacy in precisely recognizing and characterizing nanoparticle boundaries.



(a) Run 1 (b) Run 2

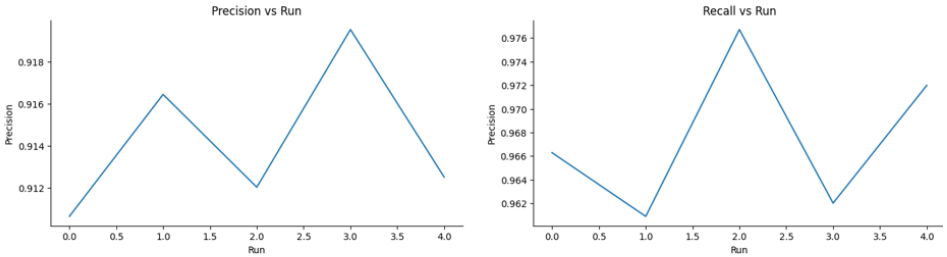


(c) Run 3 (d) Run 4



(e) Run 5

Fig. 7. Prediction outputs.



(a) (b)

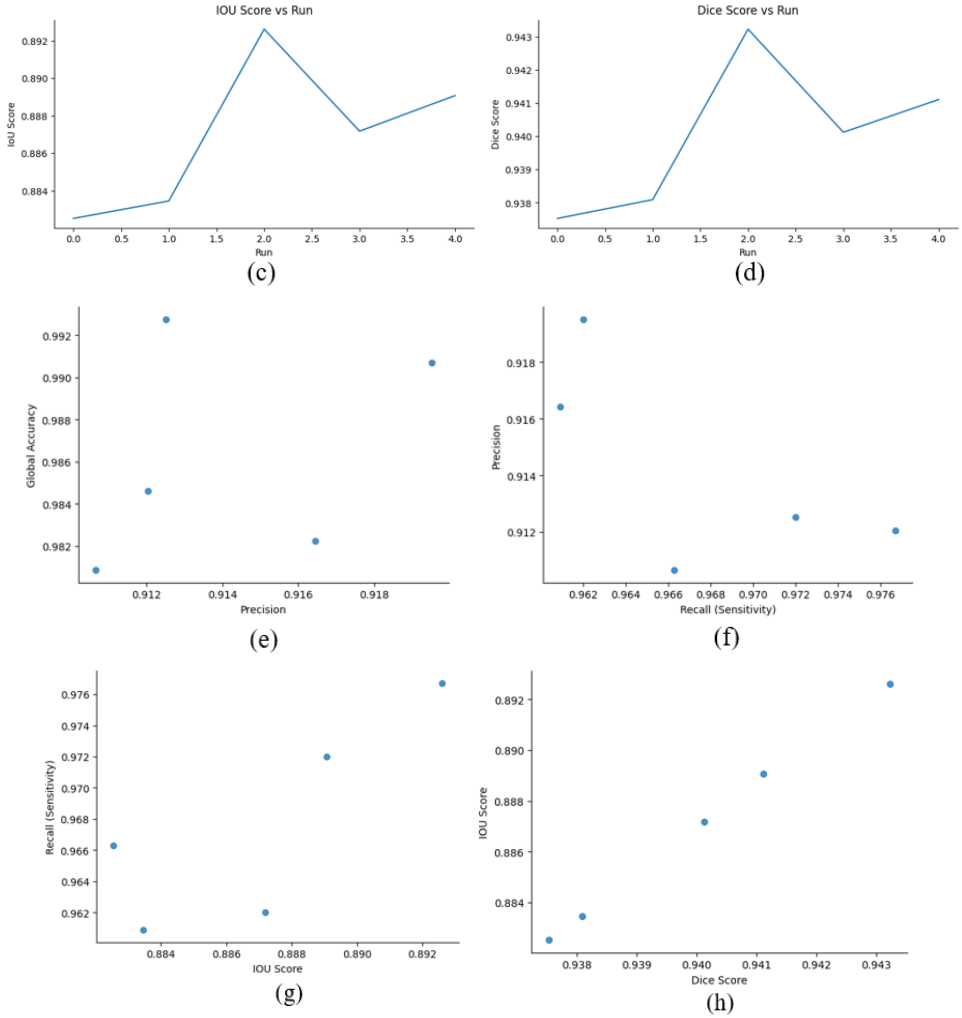


Fig. 8. Visualization of performance evaluation.

TABLE II. PERFORMANCE EVALUATION

Metric	Run 1	Run 2	Run 3	Run 4	Run 5
Dice Score (DSC)	0.939	0.938	0.938	0.943	0.946
Intersection over Union (IoU)	0.885	0.883	0.883	0.892	0.897
Recall	0.974	0.969	0.979	0.972	0.966
Precision	0.907	0.909	0.900	0.916	0.926
Global Accuracy	0.981	0.982	0.983	0.991	0.993
AUC ROC	0.990	0.992	0.988	0.994	0.994

Fig. 9 displays histograms that depict the performance measures of the mUNet model. The precision scores, between 0.912 and 0.916, demonstrate the model's robust ability to effectively separate nanoparticles with minimal false positives. Despite a slight variation in these numbers, the general stability demonstrates the model's efficacy in positive predictive accuracy. The recall scores vary from 0.962 to 0.970, indicating the model's exceptional capability in accurately identifying almost all true positive cases of nanoparticles. The IoU scores, ranging from 0.884 to 0.892, indicate strong performance, reflecting a significant overlap between expected and actual areas. The uniformity of these scores highlights the model's dependability in sustaining high-quality segmentation across samples. Finally, the Dice scores vary from 0.938 to 0.943, underscoring the model's proficiency in generating precise segmentations. The close proximity of these scores signifies negligible performance variability, hence affirming the model's dependability.

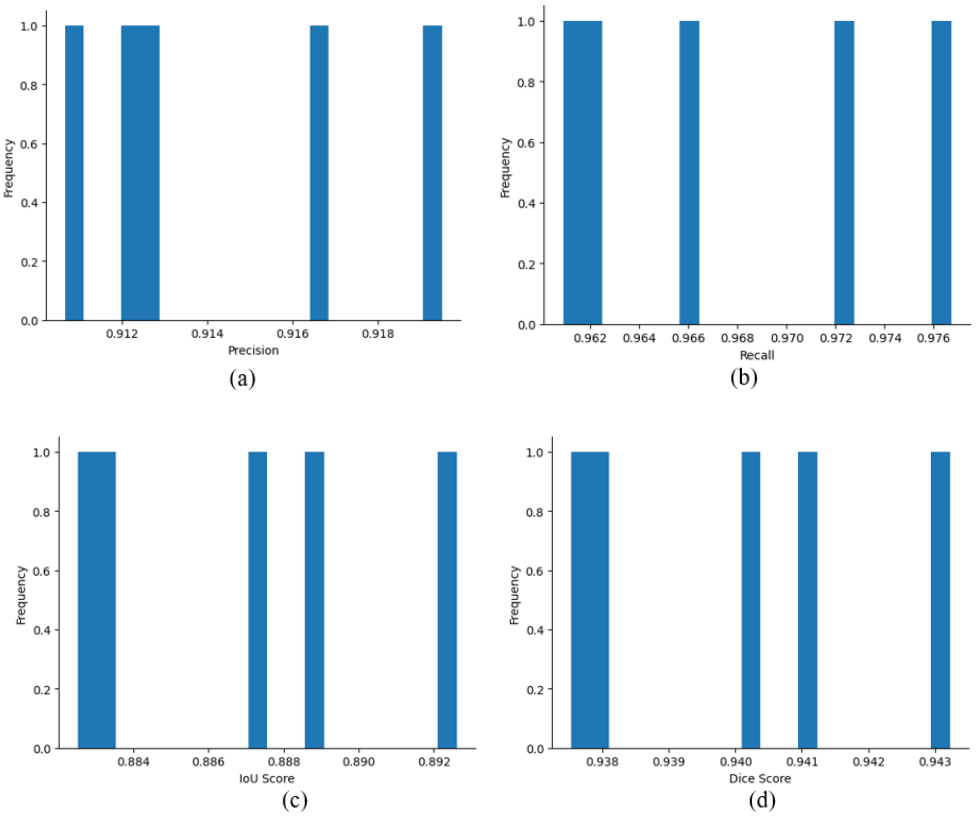


Fig. 9. Histogram plots.

Fig. 10 illustrates the mean values of the primary performance metrics acquired over an epoch of 200. The Dice Score Mean is 0.940, indicating a high degree of accuracy in the overlap between the anticipated and actual nanoparticle areas. The mean IoU score is 0.887, demonstrating strong efficacy in accurately defining nanoparticle regions with few deviations. The mean results together underscore the dependability of the mUNet model in nanoparticle segmentation tasks, demonstrating consistent performance across various parameters.

Fig. 11 depicts the standard deviations computed across 200 epochs. The Dice score demonstrates a minimal standard deviation of 0.002071, indicating consistent segmentation accuracy over repetitions. The IoU score demonstrates a standard deviation of 0.003696, signifying reliable overlap between predicted and real nanoparticle regions. The recall has slightly larger variance, with a standard deviation of 0.006005, indicating modest fluctuations in the model's capacity to recognize all actual nanoparticle occurrences.

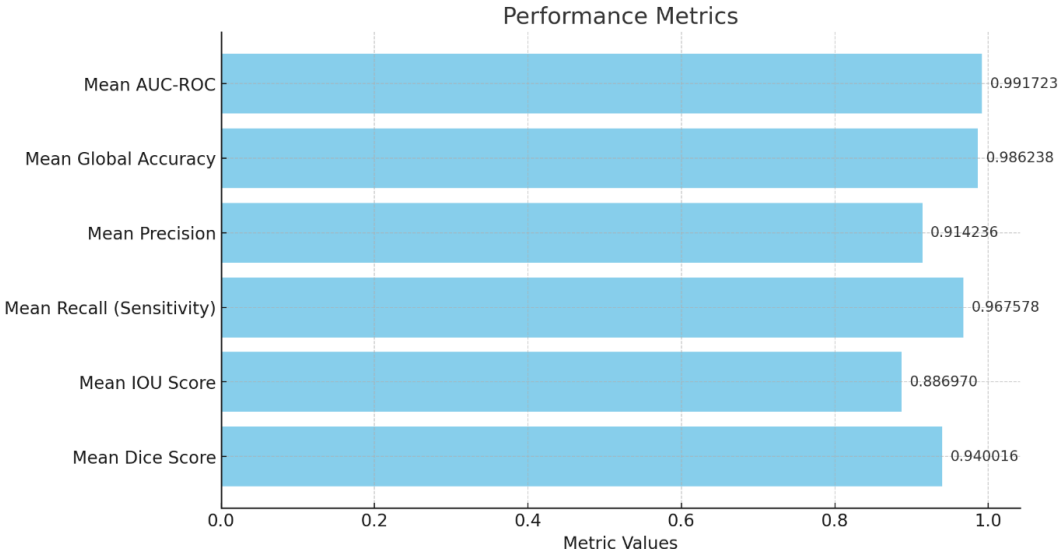


Fig. 10. Mean value of the scores.

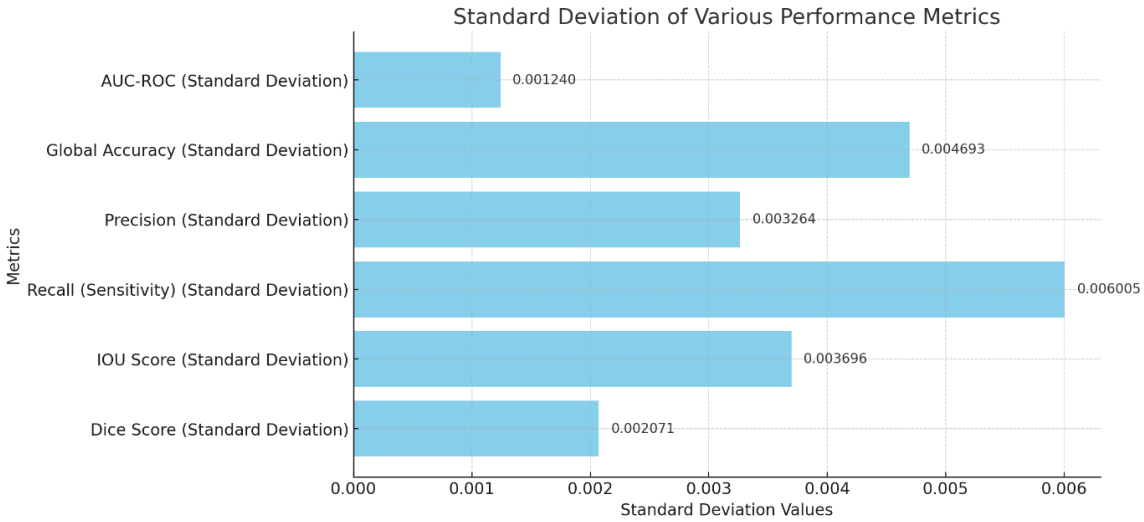


Fig. 11. Standard deviation of the scores.

The precision exhibits a standard deviation of 0.003264, signifying dependable identification of true positives. The standard deviation of global accuracy is 0.004693, indicating uniform performance across all iterations, whilst the AUC-ROC score exhibits the lowest deviation at 0.00124, implying remarkable stability in differentiating between nanoparticle and non-nanoparticle regions. The low standard deviations across all measures indicate that the mUNet model exhibits consistent and dependable performance, with minor variations in its ability to effectively segment TiO_2 nanoparticles throughout numerous iterations.

V. Discussion

A thorough accuracy comparison of the suggested model against a number of cutting-edge segmentation methods is depicted in Fig. 12 and Table III. The suggested framework attained an outstanding accuracy of 99.3%, markedly surpassing conventional approaches such as NSNet (86.2%) and more sophisticated architectures like Deeplabv3+ with ResNet-18 (96.12%), multiple-output convolutional neural networks (96.59%), and U-Net (97.1%).

TABLE III. Comparative Analysis of the Proposed Model Against Existing Methods

Author	Methodology	Accuracy (%)
Sun et al. [25]	NSNet	86.2
Gumbiowski et al. [10]	Deeplabv3+ network with a Resnet-18	96.12
Oktay et al. [24]	multiple output convolutional neural networks	96.59
Bals et al. [11]	UNET and UNet++	97
Leonid Mill et al. [23]	U-Net	97.1
Proposed model: Modified U-Net with ResNet 50		**99.3**

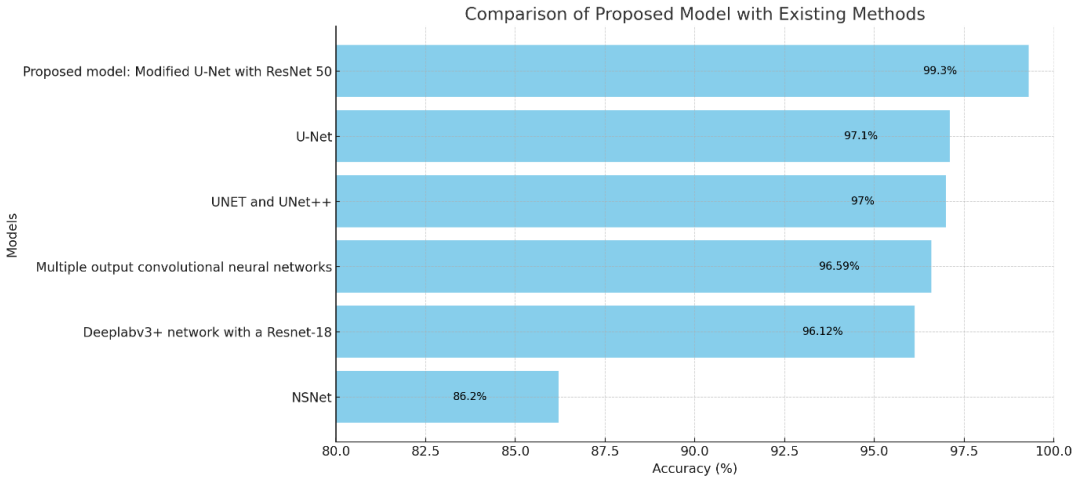


Fig. 12. Comparison of the proposed model with existing methods.

The suggested model exhibits enhanced segmentation performance compared to closely comparable models such as UNet++ (97%) and U-Net (97.1%). The enhancement is due to the incorporation of ResNet50 as the encoder, facilitating superior feature extraction via deep residual learning, along with the U-Net's strong skip connections, which ensure accurate reconstruction of spatial information. The modifications, together with the model's capacity to utilize pre-trained weights for more effective training, lead to a significant improvement in segmentation accuracy for nanoparticle boundaries in SEM images. The modified layout exhibits enhanced proficiency in managing complex textures and diverse sizes of nanoparticle areas, positioning itself as a dependable and effective solution for high-precision segmentation tasks in nanotechnology.

VI. Conclusion

The proposed research offers an extensive approach for automating the segmentation of SEM images of TiO_2 nanoparticles utilizing a modified U-Net architecture with a ResNet50 backbone pre-trained on ImageNet. This model effectively overcomes the drawbacks of conventional segmentation techniques, which frequently encounter issues with the intricate and diverse textures present in SEM images. By utilizing the robust feature extraction capabilities of ResNet50 and integrating them with the effective segmentation framework of U-Net, the model exhibits notable enhancements in accuracy, precision, and overall performance. The findings, featuring an average Dice score of 0.940 and an IoU of 0.887, demonstrate the model's proficiency in precisely delineating nanoparticle boundaries. The minimal standard deviations in all performance metrics, such as a Dice score standard deviation of 0.002071 and an IoU standard deviation of 0.003696, underscore the model's consistency and stability throughout various iterations. The incorporation of skip connections and multi-scale feature learning allows the model to preserve spatial details while analyzing abstract, high-level data. This method substantially surpasses conventional strategies in nanoparticle segmentation for accuracy and reliability, as evidenced by comparisons with existing methods. The model's versatility and diminished training duration underscore its practical utility for high-throughput nanoparticle investigation in materials science. The research highlights the efficacy of deep learning models, particularly the mUNet, in enhancing nanomaterial analysis through accurate, automated segmentation methods.

References

[1] Rokunuzzaman, M. K. (2024). The Nanotech Revolution: Advancements in Materials and Medical Science. Journal of Advancements in Material Engineering, 9(2), 1-10.

[2] Fernandez-Garcia, M., Martinez-Arias, A., Hanson, J. C., & Rodriguez, J. A. (2004). Nanostructured oxides in chemistry: characterization and properties. Chemical reviews, 104(9), 4063-4104.

[3] Shariati, L., Esmaeili, Y., Rahimmanesh, I., Babolmorad, S., Ziaei, G., Hasan, A., ... & Makvandi, P. (2023). Nanobased platform advances in cardiovascular diseases: Early diagnosis, imaging, treatment, and tissue engineering. Environmental Research, 116933.

[4] Su, D. (2017). Advanced electron microscopy characterization of nanomaterials for catalysis. Green Energy & Environment, 2(2), 70-83.

[5] Inkson, B. J. (2016). Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) for materials characterization. In Materials characterization using nondestructive evaluation (NDE) methods (pp. 17-43). Woodhead publishing.

[6] Mohale, G. T. M., Beukes, J. P., Kleynhans, E. L. J., Van Zyl, P. G., Bunt, J. R., Tiedt, L. R., ... & Jordaan, A. (2017). SEM image processing as an alternative method to determine chromite pre-reduction. Journal of the Southern African Institute of Mining and Metallurgy, 117(11), 1045-1052.

[7] Yao, L., & Chen, Q. (2023). Machine learning in nanomaterial electron microscopy data analysis. In Intelligent Nanotechnology (pp. 279-305). Elsevier.

[8] Yao, X., Wang, X., Wang, S. H., & Zhang, Y. D. (2022). A comprehensive survey on convolutional neural network in medical image analysis. Multimedia Tools and Applications, 81(29), 41361-41405.

[9] Eliasson, H., & Erni, R. (2024). Localization and segmentation of atomic columns in supported nanoparticles for fast scanning transmission electron microscopy. npj Computational Materials, 10(1), 168.

[10] Gumbiowski, N., Loza, K., Heggen, M., & Epple, M. (2023). Automated analysis of transmission electron micrographs of metallic nanoparticles by machine learning. Nanoscale advances, 5(8), 2318-2326.

[11] Bals, J., & Epple, M. (2023). Deep learning for automated size and shape analysis of nanoparticles in scanning electron microscopy. RSC advances, 13(5), 2795-2802.

[12] Larsen, M. H. L., Lomholdt, W. B., Valencia, C. N., Hansen, T. W., & Schiøtz, J. (2023). Quantifying noise limitations of neural network segmentations in high-resolution transmission electron microscopy. Ultramicroscopy, 253, 113803.

[13] Fu, W., Spurgeon, S. R., Wang, C., Shao, Y., Wang, W., & Peles, A. (2022). Deep-learning-based prediction of nanoparticle phase transitions during in situ transmission electron microscopy. arXiv preprint arXiv:2205.11407.

[14] Ji, Z., & Wang, Y. (2022). Application of Multiple Random Forest Algorithm in Image Segmentation of Nanoparticles. Journal of Nanomaterials, 2022(1), 4964368.

[15] De Backer, A., Van Aert, S., Faes, C., Arslan Irmak, E., Nellist, P. D., & Jones, L. (2022). Experimental reconstructions of 3D atomic structures from electron microscopy images using a Bayesian genetic algorithm. npj Computational Materials, 8(1), 216.

[16] Singh, H. M., Sharma, D. P., & Alade, I. O. (2022). GBR-GSO based machine learning predictive model for estimating density of Al2N3, Si3N4, and TiN nanoparticles suspended in ethylene glycol nanofluids. The European Physical Journal Plus, 137(5), 587.

[17] Okunev, A. G., Mashukov, M. Y., Nartova, A. V., & Matveev, A. V. (2020). Nanoparticle recognition on scanning probe microscopy images using computer vision and deep learning. Nanomaterials, 10(7), 1285.

[18] Horwath, J. P., Zakharov, D. N., Megret, R., & Stach, E. A. (2019). Understanding important features of deep learning models for transmission electron microscopy image segmentation. arXiv preprint arXiv:1912.06077.

[19] Wang, Y. C., Slater, T. J., Leteba, G. M., Roseman, A. M., Race, C. P., Young, N. P., ... & Haigh, S. J. (2019). Imaging three-dimensional elemental inhomogeneity in Pt–Ni nanoparticles using spectroscopic single particle reconstruction. Nano Letters, 19(2), 732-738.

[20] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.

[21] Wang, H., Cao, P., Wang, J., & Zaiane, O. R. (2022, June). Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with transformer. In Proceedings of the AAAI conference on artificial intelligence (Vol. 36, No. 3, pp. 2441-2449).

[22] Prabakaran, J., & Selvaraj, P. (2022, December). Implementation of ResNet-50 with the Skip Connection Principle in Transfer Learning Models for Lung Disease Prediction. In International Conference on

Intelligent Systems and Sustainable Computing (pp. 9-19). Singapore: Springer Nature Singapore.

[23] Mill, L., Wolff, D., Gerrits, N., Philipp, P., Kling, L., Vollnhals, F., ... & Christiansen, S. (2021). Synthetic image rendering solves annotation problem in deep learning nanoparticle segmentation. Small Methods, 5(7), 2100223.

[24] Oktay, A. B., & Gurses, A. (2019). Automatic detection, localization and segmentation of nano-particles with deep learning in microscopy images. Micron, 120, 113-119.

[25] Sun, Z., Shi, J., Wang, J., Jiang, M., Wang, Z., Bai, X., & Wang, X. (2022). A deep learning-based framework for automatic analysis of the nanoparticle morphology in SEM/TEM images. Nanoscale, 14(30), 10761-10772.