Optimizing TEM Image Segmentation: Advancements in DRU-Net Architecture with Dense Residual Connections and Attention Mechanisms

M. Nagaraju Naik
Department of ECE, CMR College of Engineering & Technology, Hyderabad, Telangana, India
nagarajunaik1976@gmail.com

Nagajyothi Dimmita
Department of ECE, Vardhaman College of Engineering, Shamshabad, Hyderabad, Telangana, India
nagajyothi1998@gmail.com

Vijayalakshmi Chintamaneni
Department of ECE, Vignan Institute of Technology and Science, Hyderabad, Telangana, India
vijji.lnctphd@gmail.com

P. Srinivasa Rao
Department of ECE, CVR College of Engineering, Hyderabad, Telangana, India
psrao.cvr@gmail.com

Nagalingam Rajeswaran
Department of EEE, Malla Reddy College of Engineering, Hyderabad, Telangana, India
rajeswarann@gmail.com (corresponding author)

Amar Y. Jaffar
Computer and Network Engineering Department, College of Computing, Umm Al-Qura University, Makkah, 21955, Saudi Arabia
ayjaafar@uqu.edu.sa

Fahd M. Aldosari
Computer and Network Engineering Department, College of Computing, Umm Al-Qura University, Makkah, 21955, Saudi Arabia
fmdosari@uqu.edu.sa

Wesam N. Eid
Cyber Security Department, College of Computing, Umm Al-Qura University, Makkah, 21955, Saudi Arabia
wneid@uqu.edu.sa

Ayman A. Alharbi
Computer and Network Engineering Department, College of Computing, Umm Al-Qura University, Makkah, 21955, Saudi Arabia
aarharbi@uqu.edu.sa

Received: 1 June 2024 | Revised: 18 June 2024 | Accepted: 20 June 2024

 DOI: https://doi.org/10.48084/etasr.7994

ABSTRACT

This study introduces an innovative enhancement to the U-Net architecture, termed Modified DRU-Net, aiming to improve the segmentation of cell images in Transmission Electron Microscopy (TEM). Traditional U-Net models, while effective, often struggle to capture fine-grained details and preserve contextual information critical for accurate biomedical image segmentation. To overcome these challenges, Modified DRU-Net integrates dense residual connections and attention mechanisms into the U-Net framework. Dense connections enhance gradient flow and feature reuse, while residual connections mitigate the vanishing gradient problem, facilitating better model training. Attention blocks in the up-sampling path selectively focus on relevant features, boosting segmentation accuracy. Additionally, a combined loss function, merging focal loss and dice loss, addresses class imbalance and improves segmentation performance. Experimental results demonstrate that Modified DRU-Net significantly enhances performance metrics, underscoring its effectiveness in achieving detailed and accurate cell image segmentation in TEM images.

Keywords-U-Net; DRU-NET; dense connections; attention blocks; SEM images

I. INTRODUCTION

Cell image segmentation is a critical process in biomedical research and clinical diagnostics, providing insights into cellular structures and their functions [1]. Accurate segmentation of cell images allows detailed analysis of cell morphology, which is essential to understanding various biological processes, diagnosing diseases, and developing targeted therapies. The segmentation process involves partitioning an image into meaningful regions, typically isolating individual cells or cellular components from the background [2]. This task is particularly challenging because of the complexity of cell structures, variations in cell shapes and sizes, and the presence of noise and artifacts in the images. Transmission Electron Microscopy (TEM) has been a transformative technology in the field of cellular imaging [3]. TEM offers unparalleled resolution, enabling researchers to visualize cellular structures at the nanometer scale. This high-resolution imaging is crucial for studying the ultrastructure of cells, including organelles, membranes, and macromolecular complexes. TEM images provide a wealth of information on cellular architecture and have been instrumental in advancing our understanding of cell biology, pathology, and the mechanisms of various diseases. However, the detailed and complex nature of TEM images also poses significant challenges for segmentation, necessitating the development of advanced image-processing techniques.

The U-Net architecture has emerged as a powerful tool for biomedical image segmentation. U-Net is a type of Convolutional Neural Network (CNN) designed specifically for segmentation [4]. It features a symmetric encoder-decoder structure, where the encoder path captures the context of the image, and the decoder path enables precise localization. Skip connections between corresponding layers of the encoder and decoder paths help preserve spatial information, which is crucial for accurate segmentation. U-Net has demonstrated remarkable performance in various biomedical imaging tasks, including cell segmentation, due to its ability to learn from relatively small datasets and its flexibility in handling different imaging modalities. Despite its success, the U-Net architecture has certain limitations as it can struggle to capture fine-grained details and maintain contextual information over long distances, which are essential for the accurate segmentation of complex TEM images. Additionally, traditional U-Net may face challenges in training convergence and feature reuse, leading to suboptimal segmentation performance.

To address these challenges, this study proposes a modification of the DRU-Net architecture, termed Modified DRU-Net, that integrates dense residual connections and attention mechanisms to enhance its segmentation capabilities. Dense connections, inspired by DenseNet architectures, allow for improved gradient flow and feature reuse by connecting each layer to every other layer in a feed-forward fashion. This dense connectivity promotes the learning of robust features and mitigates the problem of vanishing gradients, facilitating better training convergence. Residual connections are incorporated to further address the vanishing gradient issue and enable the learning of residual functions, which has been shown to improve the training of deep neural networks. These residual connections help maintain identity mappings and ensure that gradient information flows directly through the network, enhancing the overall learning process.

In addition to dense and residual connections, the proposed architecture introduces attention blocks into the upsampling path of the DRU-Net. Attention mechanisms have gained popularity in various deep-learning applications due to their ability to focus on the most relevant parts of the input. In the context of image segmentation, attention blocks help the network selectively emphasize important features and suppress irrelevant ones. By incorporating attention mechanisms, Modified DRU-Net can effectively capture fine-grained details and contextual information necessary for accurate cell image segmentation. To further enhance the training process, a poly-learning rate policy was employed, where the learning rate is dynamically adjusted based on the training progress. This approach helps in achieving better convergence and improving the model's performance. Additionally, L2 regularization is applied to reduce overfitting and ensure that the model generalizes well to unseen data. To address the issue of class imbalance, which is common in biomedical image segmentation, a combined loss function is used to combine focal and dice loss. Focal loss focuses more on hard-to-classify

examples, while dice loss ensures that the segmentation is accurate by maximizing the overlap between the predicted and ground truth masks. Combining these loss functions provides a robust optimization framework for training the Modified DRU-Net. Figure 1, shows the experimental method followed.

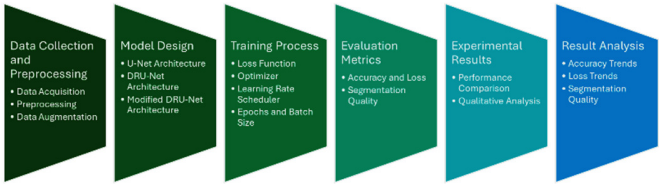

Fig. 1. Workflow of the MDRU-Net validation.

II. LITERATURE SURVEY

Traditional image segmentation techniques before the advent of deep learning included thresholding, edge detection, region-based methods, and clustering algorithms [5]. Clustering methods and contour extraction were also commonly used in the early segmentation of biomedical images [6]. The U-Net architecture includes an encoder-decoder structure with skip connections, allowing for precise localization and context integration. Various enhancements have been proposed, such as U-Net++, which introduces nested and dense skip connections to further improve segmentation performance [7, 8]. Attention U-Net incorporates attention mechanisms to focus on relevant parts of the image, improving segmentation results [9]. Comparative studies show that variants such as U-Net++ and Attention U-Net offer improved performance in terms of accuracy and robustness, particularly for complex biomedical images [10]. SD-UNet is a lightweight variant designed for low-resource environments, maintaining high accuracy with reduced computational requirements [11]. Several studies have successfully applied U-Net and its variants to cell image segmentation. The U-Net_dc model significantly improves the segmentation accuracy of endometrial cancer cells, demonstrating the model's effectiveness in complex biomedical tasks [13].

Common challenges in cell image segmentation include varying cell shapes, overlapping cells, and low-contrast images. Data augmentation, advanced pre-processing methods, attention mechanisms, and dense connections have been proposed to address these issues [14]. Variants such as Condensed U-Net (Cu-Net) tackle high-density and variable-shaped cells using improved pooling and convolution layers [15]. Different methods have been used in conjunction with U-Net for cell segmentation, including active contour models, statistical shape models, and machine learning techniques [16-19]. In [19], shape priors were used to provide robust and precise cell segmentation. Used for capturing anatomical variability, these models enhance segmentation accuracy in various biomedical applications. Performance metrics such as accuracy, precision, recall, specificity, and Intersection over Union (IoU) are commonly used to evaluate segmentation methods, and variable performance is reported based on the dataset and the segmentation technique used. Standardized datasets are essential for fair comparison and benchmarking [20].

Recent research focuses on improving U-Net with techniques such as transfer learning, reinforcement learning, and hybrid models combining multiple deep learning architectures for improved performance [21]. Combining U-Net with other architectures has shown significant improvements in segmentation accuracy and robustness [22]. U-Net has been integrated with advanced technologies to enhance its segmentation capabilities, such as Generative Adversarial Networks (GANs), and applied in multimodal imaging contexts [23-24]. Attention U-Net and its variants leverage attention mechanisms to focus on relevant image regions, improving segmentation quality [25]. Although U-Net and its variants have shown significant strengths in segmentation accuracy and efficiency, limitations include sensitivity to hyperparameters and the requirement for large annotated datasets for training, which can be resource-intensive [25-26]. Identified research gaps include the need for more robust methods to handle low-quality images, techniques to reduce the dependency on large annotated datasets, and further exploration of U-Net variants for different biomedical applications [27].

III. MATERIALS AND METHODS

This study presents an enhanced image processing model for cell image segmentation based on the DRU-Net architecture to achieve superior performance in accurately segmenting complex and variable cellular structures in TEM images. The proposed method integrates dense residual connections, attention mechanisms, multi-scale feature extraction, and boundary refinement techniques to address the limitations of existing segmentation methods. The U-Net architecture employs an encoder-decoder structure with skip connections to preserve spatial information and enhance segmentation accuracy. However, it faces challenges in capturing fine-grained details and handling geometric transformations due to its regular convolutional structure.

The proposed Modified DRU-Net incorporates deformable convolutions, Reshape-Upsampling Convolution (RUC), and focal loss. Deformable convolutions are incorporated into the encoder of the proposed Modified DRU-Net to dynamically adapt the receptive field based on the input image. This adaptation allows the network to handle geometric transformations more effectively, capturing complex shapes and structures that are often present in biomedical images. By allowing the convolutional kernels to adjust their shape and size, deformable convolutions provide the network with a greater ability to learn and represent intricate geometric variations within the data. Attention mechanisms are integrated into the network to improve its ability to focus on relevant features. These mechanisms work by selectively highlighting important areas of the input images while suppressing less significant regions. This selective focus helps the network focus on the critical features required for accurate segmentation, thus improving overall performance. Attention mechanisms are particularly beneficial in biomedical image segmentation, where the relevant structures can be small and difficult to distinguish from the background.

Dense residual connections are employed to improve the flow of gradients through the network and enhance feature

reuse. By connecting each layer to every other layer in a feed-forward fashion, they facilitate better information flow and gradient propagation during training. This design mitigates the vanishing gradient problem, allowing deeper networks to train without a significant increase in computational complexity. The result is a more robust network that can capture and utilize features more effectively. Multi-scale feature extraction is another critical improvement in the Modified DRU-Net. Biomedical images often contain structures of varying sizes, making it essential for the network to capture features at multiple scales. To achieve this, Atrous Spatial Pyramid Pooling (ASPP) is employed to allow the network to simultaneously extract features at different scales. This multi-scale approach ensures that the network can effectively capture both fine details and broader contextual information, leading to improved segmentation accuracy.

Finally, boundary-refinement techniques are integrated to enhance the accuracy of boundary segmentation. In biomedical image segmentation, accurately delineating the boundaries of structures is crucial for precise analysis. This approach includes methods such as Conditional Random Fields (CRFs) to refine the segmentation boundaries. These techniques help the network produce sharper and more accurate boundaries, reducing the likelihood of merging or splitting errors. The integration of deformable convolutions, attention mechanisms, dense residual connections, multi-scale feature extraction, and boundary refinement techniques results in a robust and efficient network capable of delivering superior segmentation performance on complex and variable cell structures in TEM images. This approach can not only improve segmentation accuracy but also enhance the network's generalization ability and computational efficiency.

The first step involves data pre-processing to ensure that the images are in a suitable format for training the model. Initially, all images were converted to grayscale, as intensity information is crucial for TEM image analysis. The pixel values were then normalized to [0, 1], standardizing the data and facilitating the training process by ensuring that the neural network operates within a consistent numerical range. Data augmentation techniques are employed to enhance the robustness and generalizability of the model, including random rotations, horizontal and vertical flipping, and elastic deformations. Data augmentation helps prevent overfitting and allows the network to learn invariant features that are not specific to the training data but generalize well to unseen images.

In the feature extraction stage, the encoder uses deformable convolutions. Unlike traditional convolutions with fixed receptive fields, deformable convolutions allow the receptive field to adapt based on the input image. This adaptive nature enables the network to capture complex geometric transformations and intricate structures that are common in TEM images. The deformable convolution layers are particularly effective in learning spatial hierarchies and relationships, making them ideal for biomedical image segmentation, where the shapes and sizes of the structures can vary significantly.

A. Model Architecture

Figure 2 shows the architecture of the proposed Modified DRU-Net, which builds upon the existing DRU-Net framework by integrating dense residual connections, attention mechanisms, deformable convolutions, multi-scale feature extraction, and boundary refinement techniques. The input layer of the model accepts grayscale images with dimensions of 512×512 pixels, normalized to the range [0, 1]

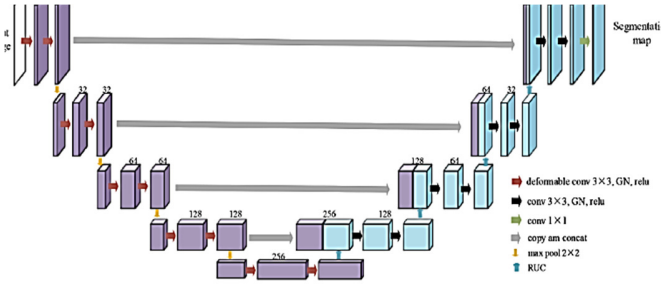

Fig. 2. Modified DRU-Net architecture

1) Encoder

- Deformable Convolutions: Each convolutional layer in the encoder is replaced with deformable convolution layers. These layers adapt the receptive field dynamically based on the input image, enhancing the model's ability to capture complex shapes and geometric variations.
- Dense Residual Connections: Each layer in the encoder is connected to every other layer in a feed-forward fashion. This dense connectivity facilitates improved gradient flow and feature reuse, addressing the vanishing gradient problem and enabling better training convergence.
- Max-Pooling Layers: Following the deformable convolutional layers, max-pooling layers reduce the spatial dimensions, allowing the network to learn hierarchical feature representations at multiple scales. The encoder comprises a series of deformable convolutional layers followed by max-pooling layers for downsampling. The standard convolution operation is given by:

$$y(p_0) = \sum_{p_n \in R} \omega p(n).x(p_0 + p_n) \quad (1)$$

where $y(p_0)$ represents the output at position x of the input image, w is the weight of the convolutional kernel, and R is the receptive field. However, in deformable convolution, the receptive field is dynamically modified as follows:

$$y(p_0) = \sum_{p_n \in R} \omega(p_n).x(p_0 + p_n + \Delta p_n) \ldots\ldots\ldots\ldots (2)$$

Here, Δp_n is a learnable offset, allowing the network to better capture local features and geometric transformations in the input images. This adaptability is crucial for accurately segmenting complex cellular structures in TEM images. The max-pooling layers are used to reduce the spatial dimensions of the feature maps, allowing the network to learn hierarchical feature representations at multiple scales. This reduction is essential for managing computational complexity and enabling the extraction of high-level features. The bottleneck layer consists of

deformable convolutions with dense residual connections to maintain high-level feature representations while mitigating information loss during downsampling.

2) Decoder

- Reshape Upsampling Convolution (RUC): The decoder employs RUC to restore the spatial resolution of feature maps while reducing the number of channels. This technique preserves spatial integrity and efficiently reconstructs high-resolution segmentation maps.
- Attention Mechanisms: Integrated into the upsampling path, attention mechanisms selectively highlight important features, enhancing the network's focus on critical areas of the input images. This improves segmentation accuracy by suppressing irrelevant information.
- Dense Residual Connections: Similar to the encoder, the decoder also uses dense residual connections to ensure efficient gradient propagation and feature reuse. The decoder employs RUC to restore the spatial resolution of feature maps while reducing the number of channels. The RUC operation is defined as:

$$RUC(x) = Reshape\big(Conv2D(x, 2C)\big) \rightarrow UpSample \rightarrow Reshape \quad (3)$$

This method avoids the use of zero-padding, thus preserving the spatial integrity of the features and efficiently reconstructing high-resolution segmentation maps from the downsampled feature representations obtained from the encoder.

- Attention Mechanisms: The upsampling path of the decoder integrates attention mechanisms to selectively focus on important features. The attention mechanism can be represented by:

$$Attention(x, g) = \sigma(Conv2D(g) + Conv2D(x)) \quad (4)$$

where σ represents the sigmoid activation function. This mechanism improves the network's ability to weigh the contributions of different feature maps, emphasizing relevant parts of the image and suppressing irrelevant ones. Attention mechanisms improve the network's ability to discriminate between different structures, leading to more accurate segmentation.

- Dense Residual Connections are employed to further enhance the flow of information and gradients through the network. A dense residual connection is given by:

$$H(x) = F(x, \{W_i\}) + x \quad (5)$$

where $H(x)$ is the output, $F(x, \{Wi\})$ is the residual mapping function with weights W_i, and x is the input. These connections facilitate better gradient propagation and feature reuse, mitigating the vanishing gradient problem and enabling the training of deeper networks. This design ensures that the network can learn more complex features and relationships, ultimately improving segmentation performance.

- Multi-scale Feature Extraction: ASPP is incorporated to capture features at different scales simultaneously. This multi-scale approach ensures that both fine details and broader contextual information are effectively captured, leading to improved segmentation accuracy. Multi-scale feature extraction techniques are employed to capture features at various scales. This is particularly important in biomedical imaging where structures vary significantly in size. ASPP is defined as:

$$y(p) = \sum_k x(p + r.k).w(k) \quad (6)$$

where r is the dilation rate that expands the receptive field without increasing the number of parameters. This allows the network to effectively capture both fine details and broader contextual information.

- Boundary refinement techniques, such as CRFs, are used to enhance the precision of boundary segmentation. The CRF model is defined as:

$$P(y|x) = \frac{1}{Z(x)} \exp\left(\sum_c \phi_c\left(y_{c,x}\right) + \sum_{c<d} \psi_{cd}\left(y_c, y_d, x\right)\right) \quad (7)$$

where $P(y \mid x)$ is the probability of label assignment given the input x, ϕ_c is the unary potential, ψ_{cd} is the pairwise potential, and $Z(x)$ is the partition function. CRFs consider the contextual dependencies between labels, leading to more accurate boundary delineation.

- Loss Functions and Learning Rate Policy: The combined loss function optimizes the model by balancing the dice loss and focal loss:

$$L_{dice} = 1 - \frac{2|X \cap Y| + 1}{|X| + |Y| + 1} \quad (8)$$

$$L_{focal} = -\alpha(1 - p_t)^{\gamma} \log(p_t) \quad (9)$$

$$L_{combined} = L_{focal} + L_{dice} \quad (10)$$

The poly-learning rate policy further stabilizes training:

$$learningrate = initiallearningrate \times \left(1 - \frac{iter}{\max\_iter}\right)^{power} \quad (11)$$

In this equation, the initial learning rate is the starting learning rate, $iter$ is the current iteration number, max_iter is the maximum number of iterations, and $power$ is a hyperparameter that controls the rate of decay. This policy helps in stabilizing the training process and achieving better convergence by gradually decreasing the learning rate as training progresses.

B. Differences between Modified DRU-Net and DRU-Net

The key differences between the proposed Modified DRU-Net and the original DRU-Net include:

- Deformable Convolutions: The Modified DRU-Net employs deformable convolutions in both the encoder and bottleneck layers, allowing adaptive receptive fields to capture complex structures, whereas DRU-Net uses standard convolutions.

- Reshape Upsampling Convolution (RUC): The Modified DRU-Net uses RUC in the decoder for efficient upsampling without zero-padding, unlike DRU-Net, which uses standard upsampling methods.
- Attention Mechanisms: Integrated into the upsampling path of the Modified DRU-Net, attention mechanisms improve focus on relevant features.
- Multi-Scale Feature Extraction: Modified DRU-Net incorporates ASPP for multi-scale feature extraction, enhancing its ability to handle features of varying sizes, a feature absent in DRU-Net.
- Boundary Refinement: Modified DRU-Net includes boundary refinement techniques such as CRFs and boundary-aware loss functions to achieve sharper and more accurate boundaries, which are not used in DRU-Net.

Together, these components create a robust and efficient network capable of delivering superior segmentation results for TEM images of cellular structures.

IV. RESULTS AND DISCUSSION

This study used the Segmentation Dataset for TEM Cell Recordings, sourced from the portal provided by the University of Freiburg [12, 28].

A. Experimental Setup and Training

The experiments were performed on a PC with an Nvidia RTX 2040 GPU. The dataset was divided into training, validation, and test sets. The images were pre-processed by normalizing the pixel values to [0, 1]. Data augmentation techniques, such as rotation, flipping, and zooming, were applied to increase the variability of the training data. The Modified DRU-Net model was trained using the Adam optimizer with an initial learning rate of 0.001. A custom poly-learning rate scheduler dynamically adjusted the learning rate during training. The training process was conducted for 50 epochs with a batch size of 8.

B. Accuracy and Loss Trends

As shown in Table I, the accuracy of the Modified DRU-Net improved significantly in the initial 10 epochs, reaching a high plateau with minimal gains beyond the 10th epoch. As shown in Figure 3, the accuracy at epoch 10 was 0.995, with a marginal increase to 0.996 by epoch 50. The training loss showed a sharp decline in the initial epochs, stabilizing after the 10th epoch. The loss at epoch 10, as shown in Figure 4, was 0.01, with a slight reduction to 0.009 at epoch 50. Validation accuracy and loss followed a similar trend, demonstrating that the model generalizes well without significant overfitting. The validation accuracy reached 0.994 by epoch 10 and 0.995 by epoch 50, while the validation loss stabilized around 0.012. U-Net started with an accuracy of 0.235 and a loss of 4.894 at epoch 1, improving to an accuracy of 0.952 and a loss of 0.4027 by epoch 10. The final accuracy at epoch 50 was 0.954. DRU-Net showed better initial performance with an accuracy of 0.435 and a loss of 2.292 at epoch 1, improving to 0.992 accuracy and 0.103 loss by epoch 10. The final accuracy at epoch 50 was 0.993, showing minimal gains beyond epoch 10.

TABLE I. COMPARATIVE ANALYSIS OF ACCURACY

Epoch	U-Net	DRU-Net	Modified MDRU-Net
1	0.235	0.435	0.256
2	0.345	0.590	0.460
3	0.475	0.735	0.610
4	0.590	0.810	0.780
5	0.695	0.875	0.850
6	0.780	0.910	0.920
7	0.845	0.940	0.960
8	0.890	0.970	0.975
9	0.920	0.985	0.985
10	0.952	0.992	0.995

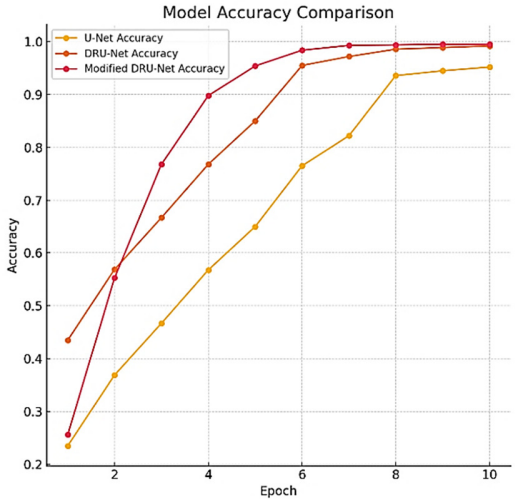

Fig. 3. Modified DRU-Net accuracy vs traditional models.

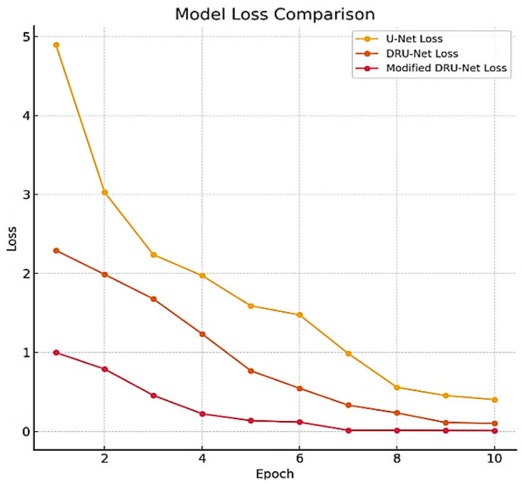

Fig. 4. Loss vs epochs Modified MDRU-Net and traditional models.

C. Segmentation Results

Figure 5 shows the segmentation results, demonstrating the qualitative performance of the models. The green outlines highlight the segmented boundaries of the cell structures in the TEM images. While the U-Net model provides a reasonable segmentation, it often struggles with fine-grained details and precise boundary delineation. The Modified DRU-Net marked improvements in capturing intricate details and maintaining contextual information, resulting in more accurate

segmentations. The Modified DRU-Net excels in both boundary accuracy and detail preservation, providing the most precise segmentation among the three models. The dense residual connections and attention mechanisms contribute significantly to these enhancements.

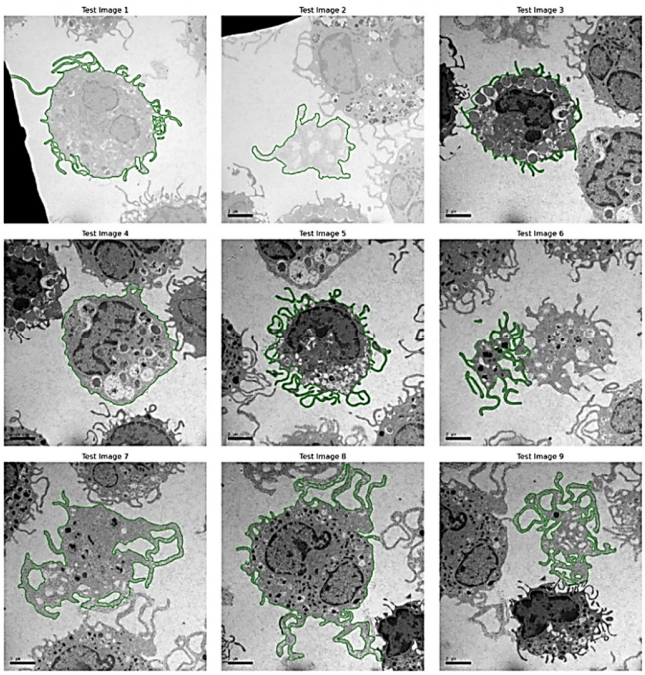

Fig. 5. Segmented test images based on the Modified DRU-Net model.

D. Detailed Analysis of Improvements

Dense residual connections facilitated better gradient flow throughout the network, ensuring effective backpropagation and mitigating the vanishing gradient problem. Attention mechanisms improved the network's ability to focus on relevant features, enhancing segmentation accuracy by suppressing irrelevant information. The deformable convolutions allowed the model to handle geometric variations and capture complex shapes more effectively, enhancing the learning and representation of intricate patterns. Multi-scale feature extraction ensured that the network could capture features at different scales simultaneously, leading to improved segmentation accuracy. The boundary refinement techniques improved the precision of segmentation boundaries, reducing errors and improving the overall accuracy.

E. Model Performance Over 50 Epochs

This study trained the proposed Modified DRU-Net model for a total of 50 epochs. The results indicate that there were no significant differences in model performance beyond the 10^{th} epoch. After the 10^{th} epoch, the accuracy and loss metrics of the Modified DRU-Net model reached a plateau, showing minimal improvement in subsequent epochs. The accuracy improvement from epoch 10 to epoch 50 was less than 0.1%, which is statistically insignificant given the overall performance gains achieved in the initial epochs. Reporting results for up to 10 epochs provides a comprehensive understanding of the model's learning curve while keeping the focus on the significant improvements made during the early stages of training. This also aligns with best practices in deep learning where early stopping is often employed to prevent overfitting and save computational resources once the model's performance stabilizes.

V. CONCLUSION

This paper demonstrates the significant advancements achieved through the development of a Modified DRU-Net architecture for the segmentation of cell images in TEM. By integrating dense residual connections and attention mechanisms into the U-Net framework, Modified DRU-Net addresses critical limitations of traditional U-Net models, such as the difficulty in capturing fine-grained details and maintaining contextual information. The experimental results highlight the superior performance of the Modified DRU-Net over both the conventional U-Net and DRU-Net models. The Modified DRU-Net achieved an impressive accuracy of 0.995 and a minimal loss of 0.01 by epoch 10, outperforming the DRU-Net and U-Net models, which reached an accuracy of 0.992 and 0.952, respectively. Rapid convergence and significant loss reduction underscore the efficiency of dense residual connections and attention mechanisms in facilitating effective gradient flow, feature reuse, and selective focus on relevant image regions. Qualitative analysis of segmentation results further validates the efficacy of Modified DRU-Net. The model consistently produced more accurate and detailed segmentations, particularly in delineating complex cellular structures compared to its counterparts. These improvements are crucial for biomedical imaging applications where precision and detail are paramount.

Although the Modified DRU-Net architecture demonstrates substantial improvements in TEM image segmentation, several areas for future research can be explored to further enhance its performance and applicability. Combining the Modified DRU-Net with advanced imaging techniques such as multimodal imaging and 3D imaging can provide a more comprehensive analysis of cellular structures, enhancing the model's utility in complex biomedical research scenarios. Additionally, integrating these techniques with real-time applications could significantly advance clinical diagnostics and personalized medicine, allowing for more precise and timely interventions. Furthermore, the use of transfer learning could facilitate the adaptation of Modified DRU-Net to various biomedical imaging modalities, reducing the need for large annotated datasets and enabling faster deployment in diverse clinical settings. Automated hyperparameter tuning and model interpretability enhancements would also contribute to more robust and user-friendly implementations, fostering greater adoption in the biomedical field. In summary, the Modified DRU-Net architecture offers a powerful tool for accurate and efficient TEM image segmentation with potential applications that extend beyond the current scope. Future research efforts that focus on broader evaluations, real-time applications, advanced imaging integration, transfer learning, automated tuning, and interpretability will continue to advance the field of biomedical image segmentation and its impact on scientific and clinical outcomes.

REFERENCES

[1] M. Gamarra, E. Zurek, H. J. Escalante, L. Hurtado, and H. San-Juan-Vergara, "Split and merge watershed: A two-step method for cell segmentation in fluorescence microscopy images," *Biomedical Signal Processing and Control*, vol. 53, Aug. 2019, Art. no. 101575, https://doi.org/10.1016/j.bspc.2019.101575.

[2] L. Heinrich *et al.*, "Whole-cell organelle segmentation in volume electron microscopy," *Nature*, vol. 599, pp. 141–146, Nov. 2021, Art. no. 7883, https://doi.org/10.1038/s41586-021-03977-3.

[3] W. Yin *et al.*, "A petascale automated imaging pipeline for mapping neuronal circuits with high-throughput transmission electron microscopy," *Nature Communications*, vol. 11, no. 1, Oct. 2020, Art. no. 4949, https://doi.org/10.1038/s41467-020-18659-3.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Munich, Germany, 2015, pp. 234–241, https://doi.org/10.1007/978-3-319-24574-4_28.

[5] B. Tian and W. Wei, "Research Overview on Edge Detection Algorithms Based on Deep Learning and Image Fusion," *Security and Communication Networks*, vol. 2022, no. 1, 2022, Art. no. 1155814, https://doi.org/10.1155/2022/1155814.

[6] S. Yin, H. Li, D. Liu, and S. Karim, "Active contour modal based on density-oriented BIRCH clustering method for medical image segmentation," *Multimedia Tools and Applications*, vol. 79, no. 41, pp. 31049–31068, Nov. 2020, https://doi.org/10.1007/s11042-020-09640-9.

[7] N. Micallef, D. Seychell, and C. J. Bajada, "Exploring the U-Net++ Model for Automatic Brain Tumor Segmentation," *IEEE Access*, vol. 9, pp. 125523–125539, 2021, https://doi.org/10.1109/ACCESS.2021.3111131.

[8] S. Das, S. Bose, Ritu, R. Jain, and M. Rout, "Light-UNet++: A Simplified U-NET++ Architecture for Multimodal Biomedical Image Segmentation," in *2023 IEEE International Conference on Contemporary Computing and Communications (InC4)*, Apr. 2023, vol. 1, pp. 1–5, https://doi.org/10.1109/InC457730.2023.10263001.

[9] P. Harsh, R. Chakraborty, S. Tripathi, and K. Sharma, "Attention U-Net Architecture for Dental Image Segmentation," in *2021 International Conference on Intelligent Technologies (CONIT)*, Hubli, India, Jun. 2021, pp. 1–5, https://doi.org/10.1109/CONIT51480.2021.9498422.

[10] R. Yousef *et al.*, "U-Net-Based Models towards Optimal MR Brain Image Segmentation," *Diagnostics*, vol. 13, no. 9, Jan. 2023, Art. no. 1624, https://doi.org/10.3390/diagnostics13091624.

[11] P. K. Gadosey *et al.*, "SD-UNet: Stripping down U-Net for Segmentation of Biomedical Images on Platforms with Low Computational Budgets," *Diagnostics*, vol. 10, no. 2, Feb. 2020, Art. no. 110, https://doi.org/10.3390/diagnostics10020110.

[12] V. Morath *et al.*, "Semi-automatic determination of cell surface areas used in systems biology," *Frontiers in Bioscience-Elite*, vol. 5, no. 2, pp. 533–545, Jan. 2013, https://doi.org/10.2741/E635.

[13] Z. Ji *et al.*, "U-Net_dc: A Novel U-Net-Based Model for Endometrial Cancer Cell Image Segmentation," *Information*, vol. 14, no. 7, Jul. 2023, Art. no. 366, https://doi.org/10.3390/info14070366.

[14] M. Ghaffari, A. Sowmya, and R. Oliver, "Automated Brain Tumour Segmentation Using Cascaded 3D Densely-Connected U-Net," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, Lima, Peru, 2021, pp. 481–491, https://doi.org/10.1007/978-3-030-72084-1_43.

[15] C. E. Akbaş and M. Kozubek, "Condensed U-Net (Cu-Net): An Improved U-Net Architecture for Cell Segmentation Powered by 4x4 Max-Pooling Layers," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Iowa City, IA, USA, Apr. 2020, pp. 446–450, https://doi.org/10.1109/ISBI45749.2020.9098351.

[16] D. L. Pham, C. Xu, and J. L. Prince, "Current Methods in Medical Image Segmentation1," *Annual Review of Biomedical Engineering*, vol. 2, pp. 315–337, Aug. 2000, https://doi.org/10.1146/annurev.bioeng.2.1.315.

[17] S. T. Acton and N. Ray, *Biomedical Image Analysis: Segmentation.* Cham, Switzerland: Springer International Publishing, 2009.

[18] Z. Linlin, H. Lu, D. Hong, and F. Huijie, "Multi-active contour cell segmentation method based on U-Net network," presented at the 12th International Conference on Image and Signal Processing (ICISP), 2020.

[19] C. Zotti, Z. Luo, A. Lalande, and P.-M. Jodoin, "Convolutional Neural Network With Shape Prior Applied to Cardiac MRI Segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 3, pp. 1119–1128, May 2019, https://doi.org/10.1109/JBHI.2018.2865450.

[20] Z. Qu, X. Tao, F. Shen, Z. Zhang, and T. Li, "Investigating Shift Equivalence of Convolutional Neural Networks in Industrial Defect Segmentation," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–17, 2023, https://doi.org/10.1109/TIM.2023.3325514.

[21] J. Zhang, Y. Xie, Y. Wang, and Y. Xia, "Inter-Slice Context Residual Learning for 3D Medical Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 40, no. 2, pp. 661–672, Oct. 2021, https://doi.org/10.1109/TMI.2020.3034995.

[22] D. Cheng and E. Y. Lam, "Transfer Learning U-Net Deep Learning for Lung Ultrasound Segmentation." arXiv, Oct. 05, 2021, https://doi.org/10.48550/arXiv.2110.02196.

[23] X. Dong *et al.*, "Automatic multiorgan segmentation in thorax CT images using U-net-GAN," *Medical Physics*, vol. 46, no. 5, pp. 2157–2168, 2019, https://doi.org/10.1002/mp.13458.

[24] Z. Lou, W. Huo, K. Le, and X. Tian, "Whole Heart Auto Segmentation of Cardiac CT Images Using U-Net Based GAN," in *2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Chengdu, China, Oct. 2020, pp. 192–196, https://doi.org/10.1109/CISP-BMEI51763.2020.9263532.

[25] N. Li and K. Ren, "Double attention U-Net for brain tumor MR image segmentation," *International Journal of Intelligent Computing and Cybernetics*, vol. 14, no. 3, pp. 467–479, Jan. 2021, https://doi.org/10.1108/IJICC-01-2021-0018.

[26] W. Shi and H. Liu, "Modified U-Net Architecture for Ischemic Stroke Lesion Segmentation and Detection," in *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, Sep. 2019, vol. 1, pp. 1068–1071, https://doi.org/10.1109/IAEAC47372.2019.8997642.

[27] H. Seo, C. Huang, M. Bassenne, R. Xiao, and L. Xing, "Modified U-Net (mU-Net) With Incorporation of Object-Dependent High Level Features for Improved Liver and Liver-Tumor Segmentation in CT Images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 5, pp. 1316–1325, Feb. 2020, https://doi.org/10.1109/TMI.2019.2948320.

[28] "TEM Dataset - Computer Vision Group, Freiburg." [Online]. Available: https://lmb.informatik.uni-freiburg.de/resources/datasets/tem.en.html.