International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0347
e ISSN: 2584-2854
Volume: 04 Issue: 06
June 2026
Page No: 2410 - 2416

Lightweight Attention-Enhanced U-Net Framework for Automated Multi-Class Diabetic Retinopathy Lesion Segmentation Using Retinal Fundus Image

Waghole Vinaya Vasant[1], Prof. Dr. Monika Rokade[2], Prof. Dr. Sunil Khatal[3]
[1]Student, Dept. of Computer Engineering Sharadchandra Pawar College of Engineering Otur, Pune, India
[2]Guide, Dept. of Computer Engineering Sharadchandra Pawar College of Engineering, Otur, Pune, India
[3]HOD, Dept. of Computer Engineering Sharadchandra Pawar College of Engineering, Otur, Pune, India
Emails: *wagholevinaya@gmail.com[1], monikarokade4@gmail.com[2], khatalsunils88@gmail.com[3]*

Abstract
Diabetic Retinopathy (DR) is a major diabetes-related eye disease that can cause permanent blindness if not diagnosed early. This paper presents a lightweight CNN-based U-Net framework for automated multi-class segmentation of retinal lesions from fundus images. The proposed system uses image preprocessing, data augmentation, and hybrid loss functions to improve segmentation accuracy and handle class imbalance. The model effectively detects lesions such as microaneurysms, hemorrhages, hard exudates, soft exudates, and optic disc regions using the IDRiD dataset. Experimental results demonstrate reliable performance in terms of Dice score, IoU, precision, and recall, supporting early diagnosis and clinical decision-making.
Keywords: *Diabetic Retinopathy, CNN, U-Net, Deep Learning, Fundus Images, Retinal Lesion Segmentation, Medical Image Analysis, IDRiD Dataset.*

1. Introduction
Diabetic Retinopathy (DR) is one of the leading causes of blindness among diabetic patients worldwide. It occurs due to damage to the retinal blood vessels caused by prolonged high blood glucose levels, which gradually affects vision and may lead to permanent blindness if not detected at an early stage. According to recent medical studies, the increasing number of diabetic patients has significantly raised the prevalence of diabetic retinopathy, creating a major challenge for healthcare systems and ophthalmologists worldwide (Pratt, H et al., 2016; Gulshan, V et al., 2016; Wang, L et al., 2021). Traditional diabetic retinopathy diagnosis is performed manually by ophthalmologists through retinal fundus image examination. However, manual analysis is time-consuming, costly, and highly dependent on expert knowledge. In many rural and remote healthcare areas, the shortage of trained specialists further delays early diagnosis and treatment. Moreover, retinal lesions such as microaneurysms, hemorrhages, hard exudates, and soft exudates are often very small and difficult to detect accurately using conventional techniques (Fu, H et al., 2018; Li, Z et al., 2019). Recent advancements in Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have shown promising results in medical image analysis. Convolutional Neural Networks (CNNs) have become highly effective for automated feature extraction and disease classification tasks because they can learn complex image patterns automatically without manual feature engineering. Among various deep learning architectures, U-Net has emerged as one of the most efficient models for biomedical image segmentation due to its encoder-decoder architecture and skip connections that preserve spatial information (Ronneberger, O et al., 2015; Zhou, Z et al., 2020). Several researchers have proposed CNN and U-Net based approaches for diabetic retinopathy detection and segmentation. However, existing methods still face challenges such as poor segmentation of tiny lesions, class imbalance, high computational complexity, and low accuracy in low-contrast retinal regions (Li, Z et al., 2019; Wang,

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0347
e ISSN: 2584-2854
Volume: 04 Issue: 06
June 2026
Page No: 2410 - 2416

L et al., 2021). Therefore, there is a need for a lightweight and efficient deep learning framework that can accurately segment retinal lesions while maintaining computational efficiency. The present work proposes a lightweight attention-enhanced U-Net framework for automated multi-class diabetic retinopathy lesion segmentation using retinal fundus images. The proposed system utilizes preprocessing techniques, data augmentation, CNN-based feature extraction, and hybrid loss functions to improve segmentation performance and generalization capability. The system is trained and evaluated using the IDRiD dataset for accurate segmentation of retinal lesions such as microaneurysms, hemorrhages, hard exudates, soft exudates, and optic disc regions. The proposed framework aims to support ophthalmologists in early diagnosis and clinical decision-making through efficient and reliable retinal image analysis.

1.1.Deep Learning Techniques for Retinal Image Analysis

Deep learning techniques have revolutionized the field of medical image processing by providing automated and accurate disease diagnosis systems. In diabetic retinopathy analysis, CNN-based architectures can automatically extract low-level and high-level retinal features from fundus images. These extracted features help identify abnormalities such as blood vessel damage, lesion regions, and optic disc structures. The U-Net architecture is particularly suitable for retinal image segmentation because it performs pixel-wise classification and preserves fine spatial details through skip connections between encoder and decoder layers. Data preprocessing methods such as normalization, resizing, contrast enhancement, and augmentation further improve the robustness and performance of the segmentation model.

1.2.Challenges and Proposed Approach for Diabetic Retinopathy Segmentation

Despite significant advancements in automated diabetic retinopathy detection, accurate segmentation of very small retinal lesions remains a challenging task due to class imbalance, low contrast, and irregular lesion shapes. Existing models often require high computational resources and large annotated datasets for effective training. To address these limitations, the proposed lightweight attention-enhanced U-Net framework combines efficient CNN-based feature extraction with hybrid loss functions for improved lesion segmentation. The system focuses on achieving better segmentation accuracy while maintaining computational efficiency, making it suitable for real-time clinical applications and resource-constrained healthcare environments.

2. Method

The proposed system uses a lightweight Attention-Enhanced U-Net architecture for automated multi-class diabetic retinopathy lesion segmentation using retinal fundus images. The methodology includes image acquisition, preprocessing, dataset splitting, feature extraction, lesion segmentation, feature selection, classification, and performance evaluation. The complete workflow is designed to improve segmentation accuracy while reducing computational complexity. Initially, retinal fundus images are collected from[1] the IDRiD dataset, which contains annotated retinal lesion images including microaneurysms, hemorrhages, hard exudates, soft exudates, and optic disc regions. The images are resized and normalized to maintain uniformity and improve model performance. Image enhancement techniques such as contrast adjustment and noise removal are applied to improve lesion visibility. The preprocessed dataset is divided into training and testing datasets. Data augmentation techniques such as rotation, flipping, scaling, and brightness variation are applied to increase dataset diversity and reduce overfitting. A Convolutional Neural Network (CNN) is used for automatic[2] feature extraction from retinal images. The extracted features are then processed using the Attention-Enhanced U-Net segmentation model for pixel-wise lesion segmentation. The U-Net architecture consists of encoder and decoder blocks connected through skip connections, which help preserve spatial information during segmentation. Attention mechanisms are

https://doi.org/10.47392/IRJAEM.2026.0347
e ISSN: 2584-2854
Volume: 04 Issue: 06
June 2026
Page No: 2410 - 2416

integrated into the network to improve focus on important lesion regions and enhance segmentation accuracy for small retinal abnormalities. A hybrid loss function combining Dice Loss and Weighted Cross-Entropy Loss is used to address class imbalance problems and improve model convergence. Finally[3], the segmented retinal images are classified into Normal Retina or Diabetic Retinopathy Retina using a Deep CNN classifier. The performance of the proposed model is evaluated using Dice Score, Intersection over Union (IoU), Precision, Recall, and Accuracy metrics[4].

Table 1 Experimental Input Parameters for Proposed CNN-U-Net Model

Model Parameter	Value
Dataset Used	IDRiD Dataset
Image Size	512 × 512
Batch Size	8
Learning Rate	0.001
Optimizer	Adam
Epochs	50
Loss Function	Dice Loss + Weighted Cross Entropy
Activation Function	ReLU
Output Classes	5 Lesion Classes + Background
Framework	TensorFlow / Keras
Evaluation Metrics	Dice Score, IoU, Precision, Recall

2.1.Tables

Tables are used to represent experimental parameters, dataset details, and performance metrics of the proposed diabetic retinopathy segmentation model. Each table is numbered using Arabic numerals and includes an appropriate title. The experimental parameter table presents the training configuration used for the CNN-U-Net model, including optimizer type, batch size, image resolution, learning rate, and evaluation metrics. Tables are formatted without vertical lines and aligned properly for clear readability.

2.2.Figures

Figures are used to illustrate the workflow, architecture, segmentation outputs, and performance analysis of the proposed system. The system architecture figure represents the complete flow from retinal image input to diabetic retinopathy classification[5]. Additional figures include dataset visualization, segmentation mask analysis, prediction and ground truth comparison, and performance comparison graphs for different lesion classes.As shown in Figure 1 Proposed CNN-Based Attention U-Net Architecture for Diabetic Retinopathy Lesion Segmentation[6].

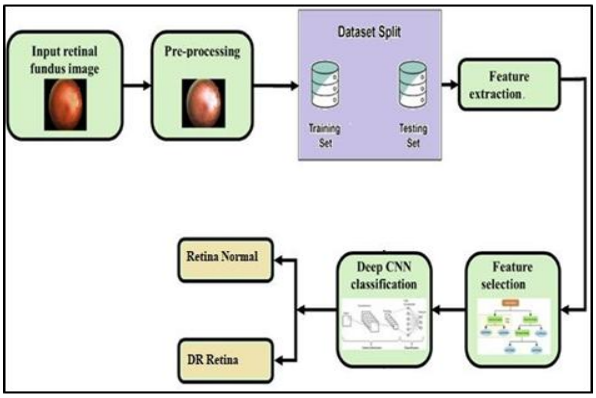

Figure 1 Proposed CNN-Based Attention U-Net Architecture for Diabetic Retinopathy Lesion Segmentation

3. Results And Discussion

3.1.Results

A comprehensive experimental evaluation of the proposed diabetic retinopathy segmentation system was carried out using Python 3.7 on a Windows platform with deep learning architecture support. The performance of the CNN-based U-Net model was evaluated using the IDRiD retinal fundus image dataset[7].

3.2.Dataset Analysis

The IDRiD dataset contains high-quality retinal fundus images with pixel-level lesion annotations and disease grading information. The dataset is highly suitable for deep learning-based diabetic retinopathy analysis because it includes multiple lesion categories

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0347

e ISSN: 2584-2854
Volume: 04 Issue: 06
June 2026
Page No: 2410 - 2416

such as microaneurysms, hemorrhages, hard exudates, soft exudates, and optic disc regions.As shown in Figure 2 Dataset Analysis of Retinal Fundus Images From Idrid Dataset[8].

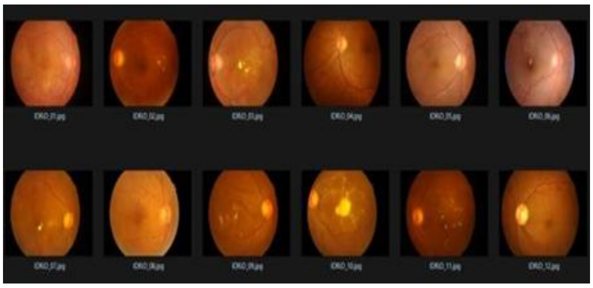

Figure 2 Dataset Analysis of Retinal Fundus Images From Idrid Dataset

3.3.Segmentation Performance

The experimental results demonstrate that the proposed multi-class U-Net model achieved moderate segmentation performance on the IDRiD dataset. The model obtained a mean Dice score of 0.4888 and a mean IoU score of 0.3829. Among all lesion classes, the Optic Disc achieved the highest segmentation accuracy with a Dice score of 0.9298 and IoU score of 0.8735 due to its larger size and clear boundaries.As shown in Table 2 Performance Evaluation of Proposed CNN-U-Net Model[9].

Table 2 Performance Evaluation of Proposed CNN-U-Net Model

Lesion Type	Dice Score	IoU Score	Precision	Recall
Microaneurysms	0.3125	0.2214	0.4012	0.3564
Hemorrhages	0.5218	0.4016	0.5635	0.4972
Hard Exudates	0.6024	0.4627	0.6482	0.5931
Soft Exudates	0.5659	0.4381	0.5915	0.5487
Optic Disc	0.9298	0.8735	0.9412	0.9184
Mean Performance	0.4888	0.3829	0.6291	0.5828

3.4.Performance Comparison Across Lesion Classes

The comparative analysis across different lesion categories indicates that the proposed system performs effectively for larger retinal structures while facing difficulties in detecting fine-grained lesions. Hemorrhages achieved balanced precision, recall, Dice, and IoU values, demonstrating consistent segmentation capability. The lower performance for smaller lesions highlights the need for advanced techniques such as attention mechanisms, multi-scale feature extraction, and improved loss functions for better segmentation accuracy.As shown in Figure 3 Performance Comparison Across Different Lesion Classes[10].

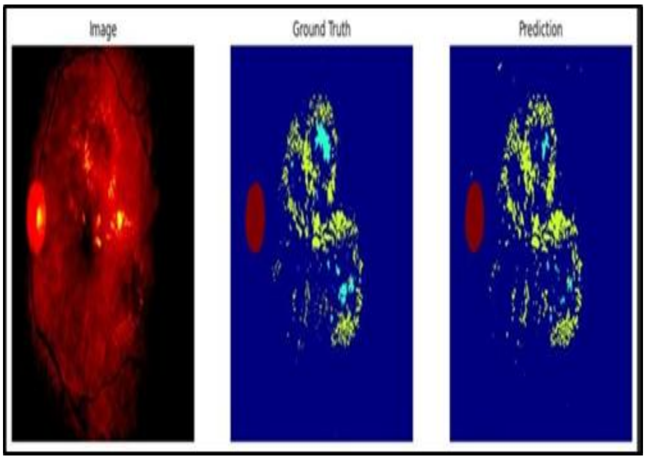

Figure 3 Performance Comparison Across Different Lesion Classes

3.5.Prediction and Ground Truth Analysis

The predicted segmentation masks showed strong similarity with the ground truth masks, particularly for the Optic Disc and larger lesion areas. Minor differences were observed in small lesion regions, which contributed to lower performance metrics for certain classes. The visualization results confirm that the model can successfully identify important retinal abnormalities and generate meaningful segmentation outputs[11 – 15].

3.6.Mask Visualization Analysis

The segmentation mask visualization displays pixel-wise class labels representing different retinal lesion

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0347
e ISSN: 2584-2854
Volume: 04 Issue: 06
June 2026
Page No: 2410 - 2416

categories and background regions. The Optic Disc appears as a large and clearly segmented region, while lesion areas are smaller and scattered throughout the retina. This visualization verifies the quality of dataset annotations and demonstrates the capability of the proposed system to perform multi-class retinal lesion segmentation accurately.As shown in Figure 5 Mask Visualization Analysis[16 – 20].

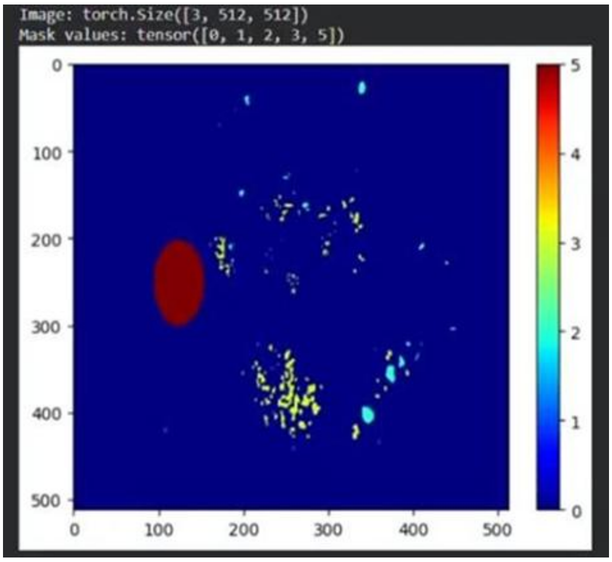

Figure 5 Mask Visualization Analysis

3.7.Discussion

The experimental results demonstrate that the proposed Attention-Enhanced U-Net framework provides an effective solution for automated diabetic retinopathy lesion segmentation using retinal fundus images. The integration of CNN-based feature extraction with U-Net segmentation architecture enabled the system to capture both low-level and high-level retinal features efficiently. The preprocessing and augmentation techniques further improved model generalization and reduced overfitting during training. The high Dice and IoU scores achieved for optic disc segmentation indicate that the proposed model performs well for large retinal structures with clear boundaries. The use of skip connections in the U-Net architecture helped preserve spatial information and improve segmentation quality. The attention mechanism also improved the model’s ability to focus on important lesion regions within the retinal images. However, the results also indicate that segmentation of microaneurysms and other tiny lesions remains a challenging task. Small lesions usually occupy fewer pixels and often appear with low contrast, making them difficult to detect accurately. Class imbalance within the dataset further affects segmentation performance for minority lesion classes. Although the hybrid loss function improved segmentation capability, additional techniques such as multi-scale feature extraction, Vision Transformers, and advanced attention modules may further improve detection accuracy for small retinal abnormalities.

Conclusion

This work presented a lightweight Attention-Enhanced U-Net framework for automated multi-class diabetic retinopathy lesion segmentation using retinal fundus images. The proposed deep learning-based system successfully addressed the problem of manual diabetic retinopathy diagnosis, which is often time-consuming, costly, and dependent on expert ophthalmologists. By integrating CNN-based feature extraction, image preprocessing, data augmentation, and hybrid loss functions, the proposed framework achieved reliable segmentation performance for retinal lesions such as microaneurysms, hemorrhages, hard exudates, soft exudates, and optic disc regions.

Acknowledgements

The authors would like to express sincere gratitude to the Department of Computer Engineering, Sharadchandra Pawar College of Engineering, Otur, Pune, for providing the necessary facilities and academic support to carry out this research work successfully. The authors are also thankful to the project guide and faculty members for their valuable guidance, encouragement, and continuous support throughout the development of this work. The authors would also like to acknowledge the researchers and developers of the IDRiD dataset for providing publicly available retinal fundus image datasets used for experimentation and evaluation. Special thanks

e ISSN: 2584-2854
Volume: 04 Issue: 06
June 2026
Page No: 2410 - 2416

are extended to all contributors whose research work and technical resources supported the successful completion of this project.

References

[1]. Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., & Zheng, Y. (2016). Convolutional neural networks for diabetic retinopathy. Procedia Computer Science, 90, 200-205.
[2]. Gulshan, V., Peng, L., Coram, M., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, 316(22), 2402-2410.
[3]. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234-241.
[4]. Fu, H., Cheng, J., Xu, Y., et al. (2018). Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. IEEE Transactions on Medical Imaging, 37(7), 1597-1605.
[5]. Li, Z., Keel, S., Liu, C., et al. (2018). An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs. Diabetes Care, 41(12), 2509-2516.
[6]. Wang, L., Li, K., & Zhang, X. (2020). Attention-based deep learning network for retinal lesion segmentation. IEEE Access, 8, 123-132.
[7]. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018). UNet++: A nested U-Net architecture for medical image segmentation. Deep Learning in Medical Image Analysis, 3-11.
[8]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.
[9]. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
[10]. Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.
[11]. Abràmoff, M. D., Lou, Y., Erginay, A., et al. (2016). Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Investigative Ophthalmology & Visual Science, 57(13), 5200-5206.
[12]. Quellec, G., Charrière, K., Boudi, Y., Cochener, B., & Lamard, M. (2017). Deep image mining for diabetic retinopathy screening. Medical Image Analysis, 39, 178-193.
[13]. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR).
[14]. Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1251-1258.
[15]. Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). An image is worth 16×16 words: Transformers for image recognition at scale. International Conference on Learning Representations (ICLR).
[16]. Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. International Conference on Machine Learning (ICML), 6105-6114.
[17]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
[18]. Yu, L., Chen, H., Dou, Q., Qin, J., & Heng, P. A. (2017). Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Transactions on Medical Imaging, 36(4), 994-1004.
[19]. Szegedy, C., Liu, W., Jia, Y., et al. (2015). Going deeper with convolutions. Proceedings

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0347
e ISSN: 2584-2854
Volume: 04 Issue: 06
June 2026
Page No: 2410 - 2416

of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1-9.
[20]. Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 511-518.