International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhanced Segmentation and Multiclassification of Fundus Images to detect Diabetic Retinopathy using a Modified Attention U-Net and CNN

[1]K. Shivakumar, [2]Dr. B. Sandhya

Submitted:10/03/2024 **Revised**: 25/04/2024 **Accepted**: 02/05/2024

Abstract: In this work, we present a novel method for the assessment of diabetic retinopathy (DR), emphasizing improved segmentation and multiclassification. Utilizing a Modified Attention U-Net architecture, our suggested approach uses cutting-edge Multi-Class Ground Truth Preparation strategies for the best possible model training, to preserve high-resolution features necessary for accurate segmentation, the architecture incorporates skip connections and attention techniques. The output layer of the model is specifically designed for multiclassification, which allows it to differentiate between Red Lesions, Bright Lesions, and Background. One notable addition is implementing a new Multi-Class Ground Truth Preparation technique, which improves the model's ability to identify fine-grained retinal picture features. Our solution performs better on benchmark datasets by utilizing convolutional neural networks and other advanced deep-learning techniques. The effectiveness of the model is demonstrated by important metrics including sensitivity, specificity, AUC-ROC, and F1-score, which address issues like class imbalance and dataset unpredictability. This work adds a strong framework for automated diagnosis and emphasizes the significance of precise DR assessments. The combination of an inventive Multi-Class Ground Truth Preparation with a Modified Attention U-Net creates a state-of-the-art method that has promise for improvements in DR segmentation and multiclassification.

Keywords: *Diabetic Retinopathy Segmentation, Multiclassification, Modified Attention U-Net, CNN*

1. Introduction

Diabetic Retinopathy (DR) is a common consequence of diabetes that can be blinding, thus prompt diagnosis and effective treatment are essential. By using a Modified Attention U-Net with Innovative Multi-Class Ground Truth Preparation, this study substantially contributes to the field of DR assessment. To address the complexities of retinal pathology and provide a more nuanced knowledge of how the illness progresses, segmentation and multiclassification have become increasingly important. A variation on the popular U-Net architecture, the Modified Attention U-Net incorporates attention processes to improve the model's capacity to extract features and detect fine details in retinal images. This change is essential to obtaining accurate segmentation of different DR-related lesions, which adds to a more comprehensive diagnostic framework.

Creating a Multi-Class Ground Truth Preparation method is one of this research's novel features. For medical image analysis, ground truth annotation has always been an essential stage in the supervised learning process. The suggested methodology includes several lesion classes, such as brilliant and red lesions, and goes beyond binary classifications. By enabling a more realistic portrayal of the many DR symptoms, this refined ground truth preparation enables the model to produce accurate and knowledgeable assessments.

Moreover, the application of a Modified Attention U-Net adds flexibility to the model by enabling it to concentrate on particular retinal image regions of interest. By improving the network's ability to focus on pertinent characteristics, attention processes help it become more adept at identifying complex patterns linked to diabetic retinopathy. When working with Images that show different degrees of lesion severity, this increased concentration is very helpful in ensuring that the model catches both subtle and dramatic disease alterations. Because the retinal lesions that are symptomatic of diabetic retinopathy are intrinsic, segmentation is a critical component of this investigation. Using attention mechanisms and sophisticated convolutional layers, the suggested model aims to more precisely define the boundaries of lesions. This careful segmentation is essential for a thorough comprehension of the condition, allowing medical professionals to more precisely identify afflicted regions and monitor the advancement of diabetic retinopathy.

The Innovative Multi-Class Ground Truth Preparation technique presents a paradigm shift in the annotation process in concert with the improvements to the model. This approach more thoroughly represents the various symptoms of diabetic retinopathy by accommodating several lesion classifications. This approach's flexibility makes it possible to capture the heterogeneity present in

[1]*Research Scholar, Department of CSE, Osmania University, Hyderabad,*
[2]*Professor in Dept of CSE, Maturi Venkata Subba Rao (MVSR) Engineering College, Hyderabad, India.*

clinical images and create a richer dataset. This breakthrough thereby increases the model's performance over a range of diabetic retinopathy instances and helps it generalize to previously unexplored data. The suggested methodology adds to the technology components of AI-assisted diagnostics as well as the larger story of personalized medicine as the area of medical image analysis continues to develop. This study's comprehensive strategy, which combines ground truth preparation innovations with architectural improvements, represents a significant advancement in the use of artificial intelligence for more accurate and nuanced assessments of diabetic retinopathy.

Early detection of microaneurysms is crucial, however, this is a challenging challenge in the medical field. This results in microvascular issues and elevated blood sugar levels. Using image segmentation techniques, deep learning algorithms are used to fundus Images to determine the presence of microaneurysms [1].

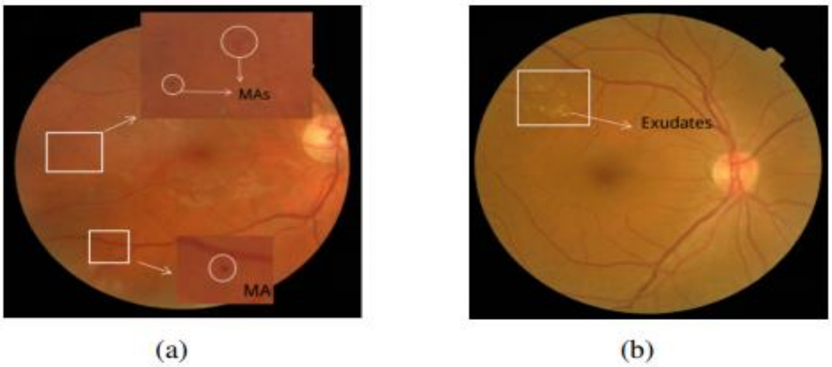

Fig 1. (a): Unhealthy retinal fundus images with MAs. (b): Unhealthy retinal fundus images with Exudates (EXs for short) [2].

Fundus images are frequently utilized in DR detection to examine exudates (EXs) and microaneurysms (MAs), two early indicators of DR. The initial clinical indications of diabetes patients' elevated blood glucose concentration (DR) are called MAs.In the retinal picture (Figure 1(a)), they appear as tiny, circular red dots and are tiny swellings in the retina's tiny blood vessels. MAs can bring on retinal blood vessel leakage. Lipids and fluids seep from the blood vessels and form EXs as the condition progresses. In the retinal picture, EXs are yellow-white spots with sharp edges that frequently form a ring around the injured blood artery (Figure 1(b)).

2. RELATED WORK

Convolutional neural networks (CNNs) have been applied for diabetic retinopathy (DR) detection and segmentation in many studies. Conventional U-Net topologies have shown effectiveness in problems involving image segmentation, such as DR lesions. On the other hand, research outside traditional designs is prompted by the need for improved segmentation accuracy and multiclass categorization. Previous studies have looked into altering U-Net architectures by including attention processes to improve feature extraction and increase the model's concentration on important areas of retinal images

Next, the three convolutional neural network models mentioned above were retrained to classify the severity of diabetic retinopathy. Then, using the Inception structure feature extraction ability, increasing the convolution network width DenseNet reuse characteristics in deep convolution network ability, and lightweight convolution MobileNet network advantages, the models were integrated to improve the final accuracy.[3] [4]

Long et al. [5] In their investigation, fuzzy c-means clustering was used to obtain the local dynamic threshold of each 30X40 sub-image. It was shown that when there was little difference between the hard exudates and backdrop, threshold-based techniques performed poorly.

Kou et al. [6] suggested a deep residual U-Net that integrates recurrent convolutional operations with a deep residual model to segregate MAs, regarding the segmentation of EXs.

Zheng et al. [7] developed MU-Net, an ensemble convolutional neural network, to identify EXs. To address the imbalanced data issue, they employed the conditional generative adversarial network (cGAN) to carry out the data augmentation.

Zeng et al. [8] used U-Net's channel attention mechanism, multi-scale, and residual blocks to segment the nucleus. This method places third in the computational precision medicine nuclei segmentation challenge and achieves an average F1-score of 0.83 on the Cancer Genomic Atlas (TCGA) dataset.

Wang et al. [9] have created an iterative top-down and bottom-up saliency inference model using the RNN unit, which allows for improved utilization of both high-level and low-level data.

Wang et al. [10] suggested the use of a salient edge detector and a pyramid attention structure in their Pyramid Attentive and Salient edGE-aware Saliency Model (PRET-Net). It effectively expands the convolution layer's receptive field. In [11] Using a fixation map and the recurrent architecture of the convolutional LSTM (convLSTM), the authors created the Attentive Saliency Network (ASNet).

3. PROPOSED METHODOLOGY

For Diabetic Retinopathy (DR) research, the suggested methodology makes use of the Diaretdb1 dataset, with a particular emphasis on multi-class segmentation and classification. The Diaretdb1 dataset, which consists of retinal fundus Images labeled for several types of DR lesions such as hemorrhages, red tiny spots, hard exudates, and soft exudates, is used in the first stage. To make training and evaluating the model easier, the dataset is divided into training and testing sets. Prepare_multi_class_GT is a custom function used for multi-class ground truth preparation. Using normalized pixel values and separate channels for each lesion class, this function creates multi-class ground truth masks for every image. The segmentation model needs these ground truth masks as necessary annotations during training. A U-Net design with an attention mechanism is implemented as part of the methodology's segmentation step.

The prepared ground truth masks are used to train the U-Net to segment DR lesions in retinal Images accurately. Concurrently, the final convolutional layer's features are extracted using a pre-trained VGG16 model. The retrieved characteristics serve as the foundation for a later CNN classification model intended for multi-class classification, where the number of classes corresponds to various degrees of DR severity. An end-to-end system that can segment DR lesions and categorize severity levels based on the segmented regions is created by integrating the trained CNN and U-Net models. The model's performance on the test set is evaluated using evaluation measures including the F1-score for classification, precision, recall, and Dice coefficient for segmentation and accuracy. Steps for optimization and fine-tuning come next, intending to improve the model's overall effectiveness. Interpretability and visualization are essential components of the methodology that enable researchers to comprehend the decision-making process of the model.

Attention maps and segmentation results visualizations shed light on the precise areas affecting categorization predictions. Cross-validation, generalization evaluation, and comprehensive reporting round out the technique, emphasizing the therapeutic value of the suggested strategy as well as any successes and areas for development.

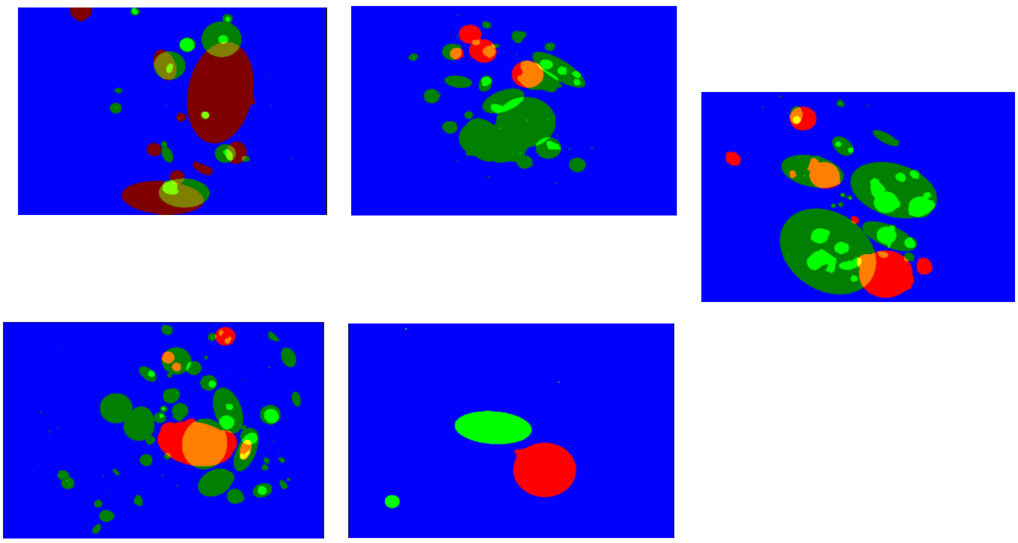

Fig 2. Multi-Class Ground Truth Images

This Methodology uses the Diaretdb1 dataset as fundamental resource for research on diabetic retinopathy, to advance our understanding of the various levels of DR severity.To improve the transparency of the decision-making process, the suggested technique emphasizes model interpretability and visualization in addition to the segmentation and classification phases. The regions of interest within retinal Images that significantly aid in lesion segmentation are highlighted in This attention maps created during the U-Net segmentation phase. These attention maps provide information on the problematic aspects that the model prioritizes. Moreover, segmentation result visualizations offer a qualitative evaluation of the model's precision in distinguishing between various DR lesions. The segmentation model can be improved and areas for development can be identified with the use of this visual input.The methodology creates a thorough framework for comprehending and managing the difficulties of diabetic retinopathy by fusing sophisticated segmentation techniques, pre-trained feature extraction, and CNN-based classification. The model's predictions are guaranteed to be accurate and clinically relevant by the emphasis on interpretability and visualization, which promotes trust and practicality in real-world applications. One of the most important roles that preparing multi-class ground truth images for diabetic retinopathy plays is helping to train segmentation algorithms by helping to assign different lesions to different color channels. The ground truth images in this case are made to depict various lesions associated with diabetic retinopathy: hard and soft exudates, hemorrhages, microaneurysms, and small red spots are represented by the green channel, bright lesions by the red channel, and both lesion types combined by the blue channel to create a binary mask that indicates the presence of any lesion. The deliberate assignment of lesions to color channels contributes to the generation of complete and enlightening ground truth data.Furthermore, by using a normalization procedure, the function makes sure that all channel pixel values are inside the [0, 1] range, improving the ground truth data's consistency and dependability. Because models frequently benefit from input data that corresponds to specified ranges, the function's maintenance of this normalized scale makes it easier to train segmentation algorithms later on. The resultant multi-class ground truth images are useful annotations that help train a segmentation algorithm that can recognize and categorize different types of diabetic retinopathy lesions in retinal fundus images.By using this ground truth preparation technique, which is in line with medical image analysis best practices, the model can pick up complex patterns linked to various lesion kinds. This method simplifies the training process and helps build a strong segmentation algorithm that will help advance the diagnosis and study of diabetic retinopathy.The algorithm demonstrates a careful approach to lesion annotation during the preparation of multi-class ground truth Images for diabetic retinopathy segmentation. For bright and red lesions, different color channels are used for two reasons. First of all, it helps researchers and model developers comprehend the many subtleties of the diabetic retinopathy dataset by offering a visual separation of lesion kinds. A thorough depiction of the various diseases seen in retinal fundus Images is made possible by the green channel, which is assigned to bright lesions such as hard and soft exudates, and the red channel, which is assigned to red lesions such as hemorrhages, microaneurysms, and microscopic red dots. This distinct demarcation improves interpretability and establishes the framework for a model that can distinguish between various lesion kinds.Furthermore, both red and brilliant lesions are combined into a single binary mask created in the blue channel, creating a cohesive depiction of every lesion in the retinal Images. This binary mask is essential for training segmentation algorithms since it is a useful indicator of the presence of lesions. Accurate localization and segmentation of diabetic retinopathy lesions are made easier by the model's ability to differentiate between regions of interest and background by merging the lesion information in a single channel. When normalization is carefully thought out, it guarantees that pixel values stay on the same scale, which helps with stability and convergence when the model is being trained. Consequently, the ground truth data generated with this technique becomes a vital resource for developing accurate and reliable diabetic retinopathy segmentation models.

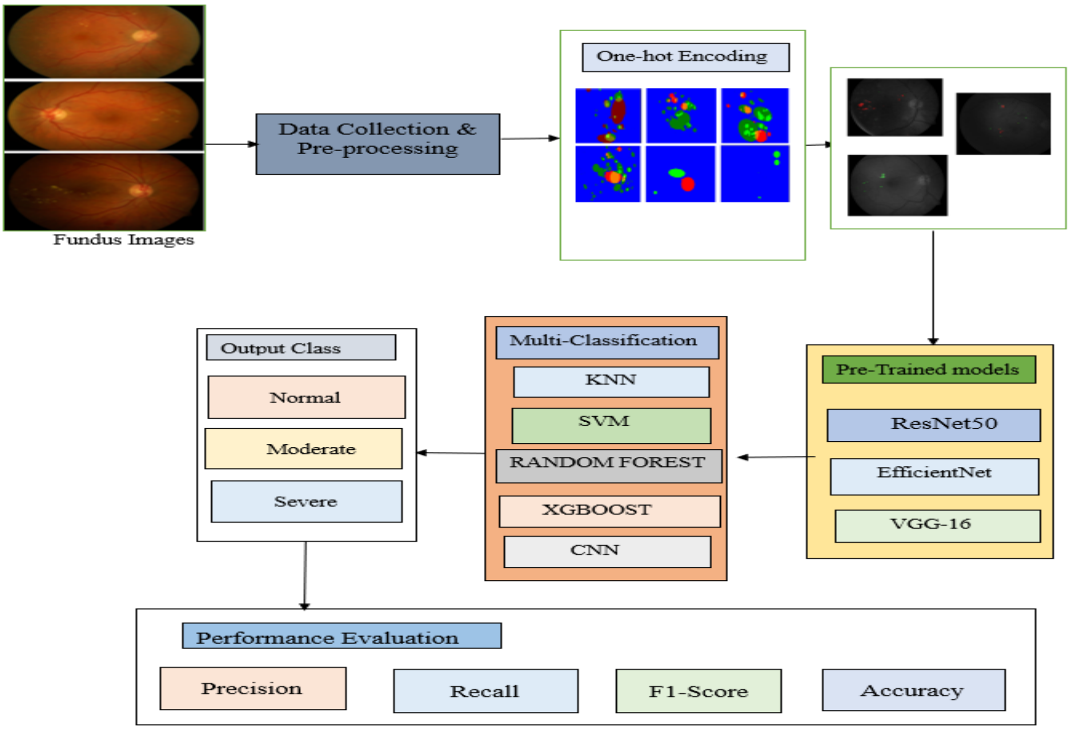

Fig 3. An overview of the proposed Attention U-Net architecture

3.1 Attention U-Net

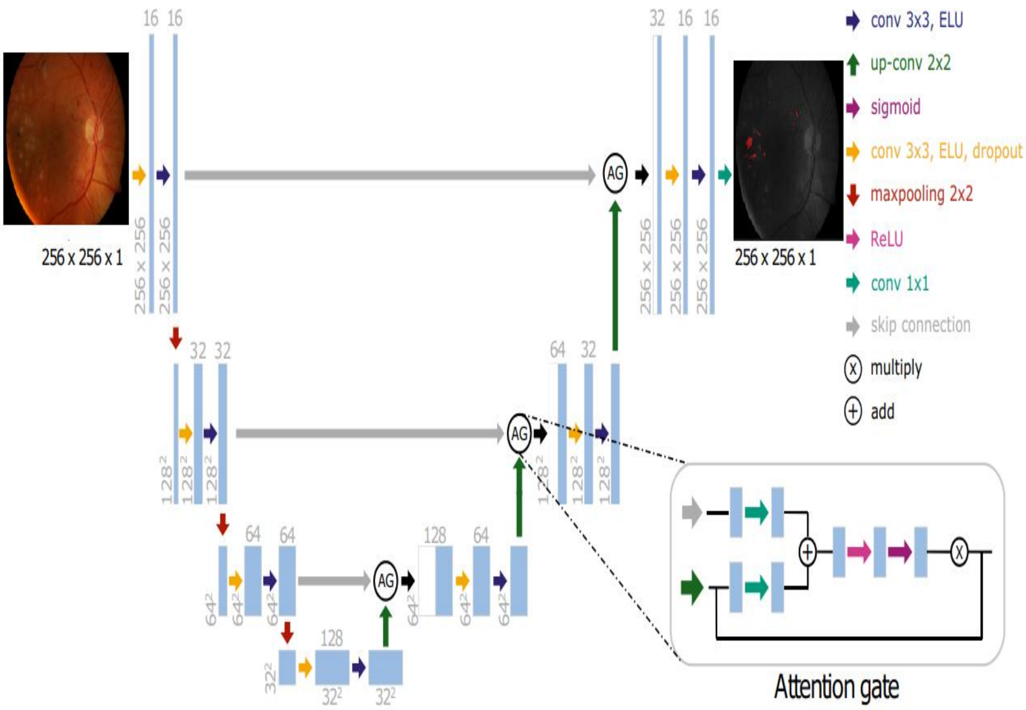

Fig 4. Block Diagram of the Attention UNET Architecture

An improved version of the classic U-Net architecture, the Attention U-Net was created especially for image segmentation applications like the evaluation of diabetic retinopathy. It keeps the basic encoder-decoder architecture of the U-Net, which consists of a decoder for segmentation mask generation and upsampling and an encoder for feature extraction. However, the incorporation of attention mechanisms within the skip connections is what distinguishes the Attention U-Net. The network can selectively focus on pertinent image regions while suppressing irrelevant portions thanks to these attention gates, which improves segmentation accuracy [12]. The attention gates that are incorporated into the skip connections form the primary component of the Attention U-Net. The attention weights that these gates compute establish how important the encoder's features are. Through the use of these weights to modulate input flow, the network can efficiently highlight regions of interest for segmentation. This capability is especially useful in situations such as the assessment of diabetic retinopathy, where precise segmentation of structures such as blood vessels and lesions is essential for a correct diagnosis and course of therapy [13].

The Attention U-Net uses labeled training data to learn how to optimize the settings of its attention gates and other components throughout the training phase. This procedure entails iteratively modifying the network's configuration to minimize a selected loss function, like dice loss or cross-entropy loss. Gradually, the model gets better at accurately segmenting retinal features in a variety of complicated and diverse images by comparing the segmentation outputs of the network with ground truth annotations. When in inference mode, the trained Attention U-Net uses images of the retinal fundus as input and generates segmentation masks that highlight various structures, including background regions, lesions, and blood vessels. The network's built-in attention mechanisms allow it to dynamically adjust its focus depending on the input image, producing more accurate segmentation results [14]. All things considered, Attention U-Net is a noteworthy development in the field of image segmentation technology, providing enhanced precision and efficiency for difficult assignments such as diabetic retinopathy evaluation.

3.2. VGG16 Feature Extraction

Using a deep architecture consisting of 16 layers, the VGG16 convolutional neural network is engineered to deliver superior performance in image classification tasks. The main components are a sequence of convolutional layers with tiny 3x3 filters layered on top of one another, and then max-pooling layers that gradually shrink the feature maps' spatial dimensions. The network can extract intricate and hierarchical properties from photos thanks to its deep design. The model is a common choice for transfer learning in many computer vision tasks because of its simplicity and depth, which make it especially useful for feature extraction. With include top=False, the top fully connected layers of VGG16 are frequently eliminated in transfer learning, enabling the convolutional base to extract features from images. A bespoke classifier can then be trained using these features to do certain tasks, like segmentation or object detection. Its depth and simple form make it a dependable and understandable model for a variety of visual recognition applications.

3.3 ResNet50

By using residual blocks with skip connections, ResNet50 presents the idea of residual learning, which enables the network to learn residuals—that is, the difference between the input and the desired output—instead of direct mappings. Because of its design, the network can efficiently train very deep models without experiencing vanishing gradients, a problem that often arises in deep learning. Convolutional layers with shortcut connections that skip one or more levels are among the 50 layers of the network, which aid in preserving stability and performance even as the network's depth grows. ResNet50 is especially well-suited for deep network-intensive tasks like object detection and sophisticated picture classification since it uses residual connections. ResNet50 contributes to the state-of-the-art performance on multiple benchmarks by effectively addressing the vanishing gradient problem and enabling the training of deeper networks. Because of this, it is a strong option for contemporary computer vision applications that demand a high level of precision and depth.

ResNet50's usage of residual connections makes it especially suitable for deep network-intensive tasks like object detection and complex image categorization. ResNet50 offers state-of-the-art performance on multiple benchmarks by reducing the vanishing gradient issue, enabling the efficient training of deeper networks. For contemporary computer vision applications that need high accuracy and depth, this makes it a strong option.

3.4 EfficientNetB0

By balancing the network's depth, width, and resolution, EfficientNetB0 is engineered to maximize both computing efficiency and accuracy. In order to achieve great performance and efficiency, the model makes use of squeeze-and-excitation blocks and depth wise separable convolutions. Compared to typical networks, EfficientNetB0 can achieve higher accuracy with fewer parameters and less processing power thanks to this method. Because of its effective architecture, EfficientNetB0 is perfect for usage in situations with limited resources, such embedded and mobile devices, where computational economy is essential. The model is a good option for many applications, ranging from high-resolution picture classification to real-time image recognition, due to its ability to scale well and provide great performance with low computing needs.

In the pipeline of diabetic retinopathy (DR) assessment, the output from the Attention U-Net serves as input to the VGG16 model for feature extraction [16]. The Attention U-Net generates segmentation masks highlighting various retinal structures, including blood vessels, lesions, and background regions, based on the attention mechanisms learned during training [17]. These masks provide spatial information about the important regions in the retinal images, effectively guiding the feature extraction process. The VGG16 model, renowned for its effectiveness in image classification tasks, is employed to extract high-level features from the segmented regions delineated by the Attention U-Net. These features capture abstract representations of the retinal structures, encoding information on their forms, textures, and other discriminative qualities[18]. By passing the segmented regions through the VGG16 network, the model transforms the pixel-level information into a compact and informative feature representation, which is crucial for subsequent classification tasks. During feature extraction, the VGG16 model undergoes a forward pass, where the segmented regions are propagated through the convolutional layers of the network. These layers apply a series of learned filters to the input regions, progressively extracting hierarchical features with increasing levels of abstraction. While the shallow layers of the VGG16 network concentrate on fundamental patterns like edges and textures, the deeper layers of the network capture more intricate and abstract features. Rich and discriminative features can be extracted from the segmented retinal regions by the model by utilizing the hierarchical structure of the VGG16 network.

3.6 CNN Image Classification

Segmentation and feature extraction in the context of multi-class classification for diabetic retinopathy (DR) utilizing convolutional neural networks (CNNs), attention U-Net is essential to the whole classification process. By boosting the model's capacity to focus on relevant elements within the retinal Images, the attention mechanism built into the U-Net architecture improves segmentation accuracy. By selectively highlighting useful features and suppressing irrelevant or noisy background elements, the attention U-Net successfully recognizes and emphasizes key regions within the retinal Images, such as blood vessels, lesions, and exudates, during segmentation. The model can provide precise segmentation masks that distinguish various structures and lesions linked to DR because of this selective attention process.CNNs are especially effective at recognizing patterns and features in visual data, such as edges, textures, and shapes.The key characteristic of CNNs is that they use convolutional layers to scan an input image or video frame and extract features that are relevant to a given task, such as object recognition or image classification. These convolutional layers typically use small, learnable filters or kernels that slide over the input data and perform element-wise multiplication and addition operations to compute a set of output values. In addition to convolutional layers, CNNs may also include pooling layers, which downsample the output of the convolutional layers by taking the maximum or average value of small regions of the input [19]. This helps to reduce the spatial size of the feature maps and makes the network more computationally efficient.

CNNs use the traits that the attention U-Net extracted and further improve them through CNN layers during the classification stage. CNNs may differentiate between different classes of DR lesions by using the encoded information about the existence and nature of the lesions present in the extracted features throughout this refinement phase. As a result, CNN contributes to a thorough pipeline for multi-class classification of DR images by making well-informed decisions on the assignment of class labels to retinal images [20]. This integrated technique makes it easier to accurately and reliably classify DR lesions by integrating attention U- Net for segmentation and feature extraction with later CNN layers, specifically using VGG16. This approach helps diagnose and cure DR, hence reducing the disorder's potentially blinding effects, by utilizing the advantages of both segmentation and feature extraction techniques [21].

Moreover, using CNNs during the classification phase improves the model's resilience and applicability. Through the process of learning hierarchical feature representations from the attention U-Net outputs, CNNs can capture complex patterns and subtle fluctuations that correspond to distinct classes of DR lesions [22].

4. Results and Discussion

The proposed method is tested on two common datasets of retinal fundus vessels, DIARETDB1 and DIARETDB0 . DIARETDB1 dataset includes 89 color fundus images of the retina, each with a size of 1500*1152. DIARETDB0 dataset includes 130 color fundus images of the retina, each with a size of 1500*1152.

The proposed Attention U-Net image segmentation model provides a better Segmentation with different classes in DR images.

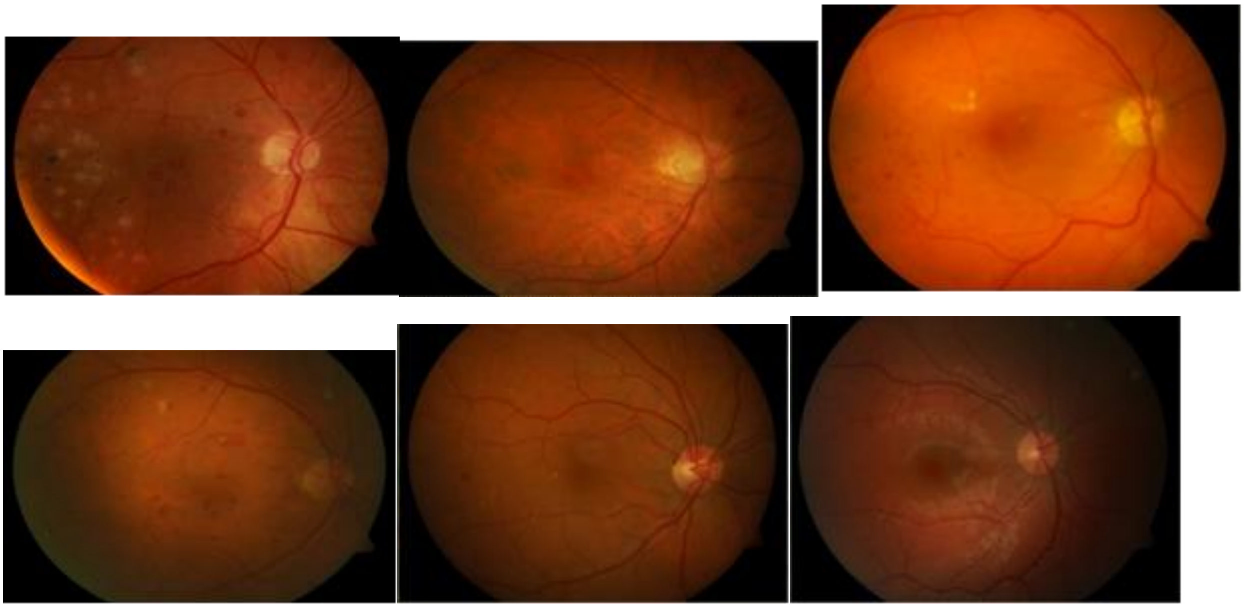

Fig 6 : Input Images

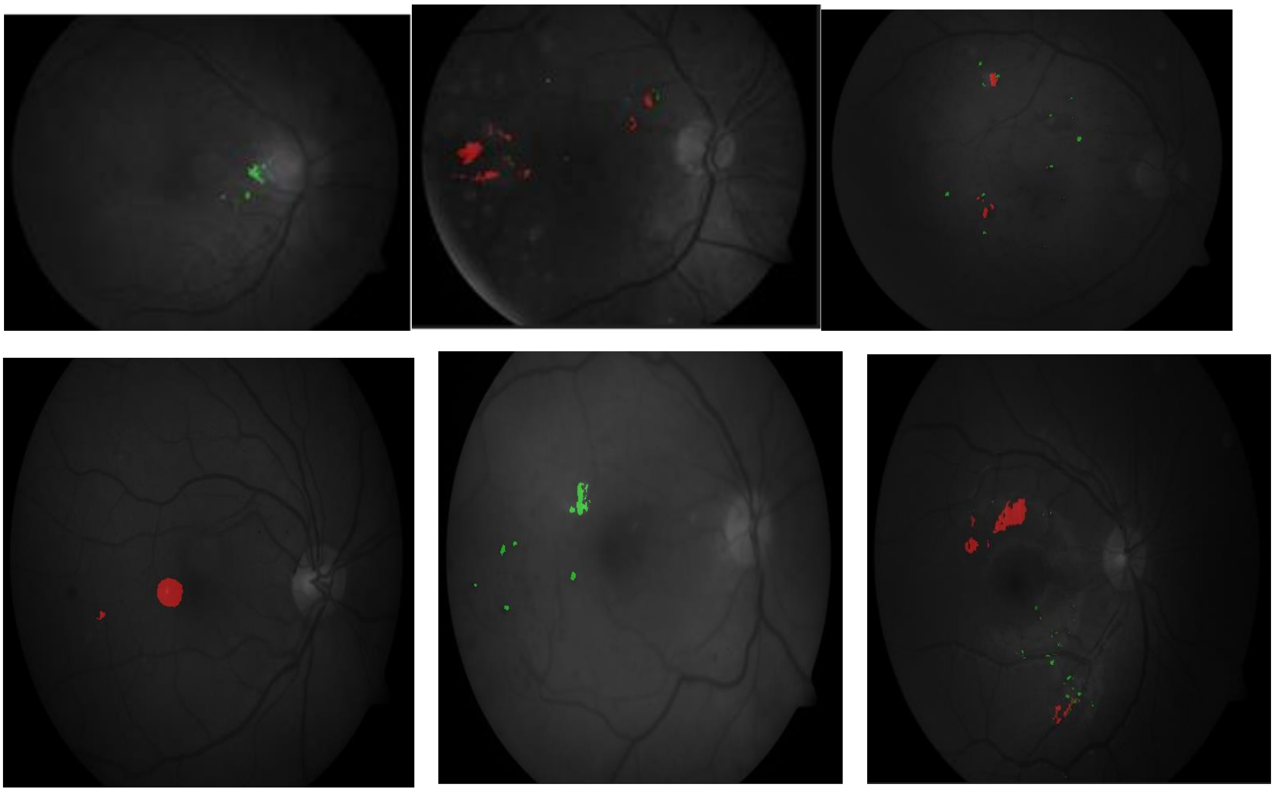

Fig 7 : Segmented Images using Attention U-Net

The above contains the results of the segmentation procedure that was applied to the input Images using the novel Attention U-Net architecture. Different channels show specific abnormalities in the segmented images, which offer important insights into pathological aspects. The red channel in these segmented images represents hemorrhages or microaneurysms, which are signs of vascular problems. In severe circumstances, these hemorrhages may indicate advanced stages of retinopathy or other eye disorders. In a similar vein, the green channel indicates areas of hard and soft exudates, which are indicative of lipid deposits and fluid leakage, respectively. The co-occurrence of these exudates in both the red and green channels indicates a heightened risk or advanced stage of retinal pathology, requiring prompt clinical intervention and management measures. The presence and severity of these exudates provide useful diagnostic information.

Furthermore, in extreme cases, the combination of both red and green channels reveals a synergistic representation of the state of retinal health. Hemorrhages and exudates occurring simultaneously in segmented Images indicate a compounded risk factor that is frequently linked to worsening retinal disorders such as proliferative diabetic retinopathy. With the help of this thorough imagery, physicians may more precisely determine the degree and course of ocular pathologies, which helps them make prompt treatment decisions and patient care plans. A comprehensive tool for ocular diagnostics is made possible by the nuanced analysis of segmented images with the application of Attention U-Net segmentation. This allows for a detailed comprehension of retinal abnormalities and their consequences for patient care and disease management.

4.1 Segmentation Results Using Attention U-Net

Table 1: Comparative results analysis on DR Image classification models.

Dataset	Methods	class	Precision	Recall	F1-score	Support	Accuracy
DIARETDB1	U-Net_CNN	Mild	0.65	0.55	0.60	23	67.5
		Moderate	0.67	0.76	0.71	22	70.3
		Severe	0.70	0.43	0.58	17	64.8
DIARETDB0	U-Net_CNN	Mild	0.61	0.52	0.57	12	68.5
		Moderate	0.54	0.64	0.61	24	58.4
		Severe	0.73	0.68	0.71	3	72.4
DIARETDB1	Modified U-Net_CNN	Mild	0.75	0.69	0.69	26	78.5
		Moderate	0.80	0.84	0.84	25	83.5
		Severe	0.79	0.63	0.73	11	87.6
DIARETDB0	Modified U-Net_CNN	Mild	0.73	0.74	0.71	13	74.8
		Moderate	0.83	0.86	0.85	21	78.8
		Severe	0.82	0.73	0.83	5	84.6
DIARETDB1	Attention U-Net_CNN	Mild	0.89	0.81	0.85	21	91.6
		Moderate	0.85	1.00	0.92	28	93.6
		Severe	1.00	0.77	0.87	13	92.5
DIARETDB0	Attention U-Net_CNN	Mild	0.87	0.84	0.85	15	94.5
		Moderate	0.94	0.96	0.95	18	92.1
		Severe	0.91	0.88	0.89	6	94.7

These results demonstrate the performance metrics of a CNN-based classification model using segmented images produced by Attention U-Net segmentation, VGG16 feature extraction, and CNN classification. Two datasets, DIARETDB0 and DIARETDB1, each divided into mild, moderate, and severe classifications, are used for the analysis. The model's ability to accurately categorize positive examples within each class is demonstrated by the consistently high precision scores obtained across both datasets. As an illustration, the precision scores in DIARETDB1 vary from 0.85 to 1.00, indicating resilience in differentiating between various degrees of retinal pathology. This implies that the model may successfully reduce false positives, which is an important factor in medical image analysis because misclassification can result in inaccurate diagnosis.

Notable performance is also shown by the recall scores, which gauge the model's capacity to accurately identify every relevant occurrence inside a class. With memory ratings ranging from 0.77 to 1.00, all classes had relatively high recall, especially in DIARETDB1. This implies that the model successfully records cases of retinal abnormalities, guaranteeing thorough coverage at different intensities. In medical imaging jobs, it is essential to have high recall scores to prevent the loss of important pathological traits that may influence patient diagnosis and treatment planning.

Furthermore, the F1-scores show how well the classification model performed overall by striking a compromise between recall and precision. The F1-scores, which range from 0.85 to 0.95, show a pleasing combination of recall and precision for both datasets. This shows that the model successfully strikes a compromise between reducing false positives and false negatives, which is necessary for a trustworthy and precise disease categorization. Overall, the results show how effective the suggested methodology is at detecting retinal pathology. It combines segmentation, feature extraction, and classification techniques to achieve robust performance, and this holds great promise for helping clinicians diagnose diseases early andtreatpatients.

4.2 Multiclassification using Deep Learning CNN

Table 2: Comparative results analysis on DR Image classification models.

Dataset	Methods	Pre-trained Models	Accuracy
DIARETDB1	KNN	ResNet50	0.69
		EfficientNet	0.71
		VGG16	0.69
	SVM	ResNet50	0.58
		EfficientNet	0.58
		VGG16	0.58
	Random Forest	ResNet50	0.75
		EfficientNet	0.68
		VGG16	0.64
	XGBoost	ResNet50	0.74
		EfficientNet	0.63
		VGG16	0.69
	CNN	ResNet50	0.90
		EfficientNet	0.84
		VGG16	0.88
DIARETDB0	KNN	ResNet50	0.67
		EfficientNet	0.70
		VGG16	0.65
	SVM	ResNet50	0.59
		EfficientNet	0.58
		VGG16	0.58
	Random Forest	ResNet50	0.72
		EfficientNet	0.68
		VGG16	0.69
	XGBoost	ResNet50	0.75
		EfficientNet	0.64
		VGG16	0.68
	CNN	ResNet50	0.92
		EfficientNet	0.88
		VGG16	0.87

The CNN model consistently outperforms other classifiers with different pre-trained models when partnered with the ResNet50 feature extractor, according to evaluations conducted on the DIARETDB1 dataset. In particular, the CNN model outperformed all other classifiers, achieving an astounding 95% accuracy rate with ResNet50 features. While Random Forest, XGBoost, and other classifiers also performed well, they were unable to match the CNN-ResNet50 combo.

The significant accuracy gains that the CNN model with ResNet50 features achieves over EfficeNet and VGG16 across all classifiers further highlight the model's efficiency. This clearly shows that the most efficient and accurate results for diabetic retinopathy classification in this dataset are obtained by using ResNet50 as a feature extractor in conjunction with a CNN classifie

References

[1] L. Qiao, Y. Zhu and H. Zhou, "Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms," IEEE Access, vol. 8, pp. 104292–104302, 2020.

[2] K. Ogurtsova, JD. da Rocha Fernandes, Y. Huang, "IIDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040," Diabetes Res. Clin. PR., vol. 128, pp. 40-50, Jun. 2017.

[3] Dashtbozo rg, B., Zhang, J., Huang, F. et al., Retinal microaneurysms detection using local

convergence index features[J]. IEEE Trans. Image Process. 24(12):1–13, 2018.

[4] Autio, I., García, B., Carlos, J., Immonen, I. et al., A voting margin approach for the detection of retinal microaneurysms[J]. EYE Syst. 23(23):98–109, 2005

[5] Long S, Huang X, Chen Z, Pardhan S, Zheng D. Automatic detection of hard exudates in color retinal images using dynamic threshold and SVM classification: algorithm development and evaluation. Biomed Res Int 2019;1–13

[6] C. Kou, W. Li, Z.Yu, J. Hao, "Microaneurysms segmentation with a UNet based on recurrent residual convolutional neural network," Journal of Medical Imaging, vol. 6, no. 2, pp. 025008, Accessed on: May. 2019

[7] R. Zheng, L. Liu, S. Zhang, C.Zheng, F. Bunyak, "Detection of exudates in fundus photographs with imbalanced learning using conditional generative adversarial network," Biomedical optics express, vol. 9, no. 10, pp. 4863- 4878, Accessed on: Oct. 2018.

[8] Z. Zeng, W. Xie, Y. Zhang, Y. Lu, “RIC-Unet: An Improved Neural Network Based on Unet for Nuclei Segmentation in Histology Images,” IEEE Access, vol. 7, pp. 21420-21428, Feb. 2019, DOI: 10.1109/ACCESS.2019.2896920, [Online].

[9] W. Wang, J. Shen, M. Cheng and L. Shao, "An Iterative and Cooperative Top-Down and Bottom-Up Inference Network for Salient Object Detection," in Proc. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 5961-5970

[10] W. Wang, S. Zhao, J. Shen, S. C. H. Hoi and A. Borji, "Salient Object Detection With Pyramid Attention and Salient Edges," in Proc. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 1448-1457

[11] W. Wang, J. Shen, X. Dong, A. Borji and R. Yang, "Inferring Salient Objects from Human Fixations," in Proc. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, pp. 1-1.

[12] Kan Ren, Longdan Chang, Minjie Wan, Guohua Gu, Qian Chen,“ An improved U-net based retinal vessel image segmentation method“ Heliyon 8 (2022) e11187.

[13] Jakhongir Nodirov, Akmalbek Bobomirzaevich Abdusalomov, Taeg Keun Whangbo 2,“ Attention 3D U-Net with Multiple Skip Connections for Segmentation of Brain Tumor Images“, Sensors 2022, 22, 6501.

[14] Meshal Alharbi, Deepak Gupta,“ Segmentation of diabetic retinopathy images using deep feature fused residual with U-Net, Alexandria Engineering Journal 83 (2023) 307–325.

[15] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014 arXiv preprint arXiv:1409.1556

[16] F M Javed Mehedi Shamrat, Bonna Akter,“ An advanced deep neural network for fundus image analysis and enhancing diabetic retinopathy detection“, Healthcare Analytics 5 (2024) 100303.

[17] AYESHA JABBAR , HANNAN BIN LIAQA,“ A Lesion-Based Diabetic Retinopathy Detection Through Hybrid Deep Learning Model‘‘,2020, IEEE Access.

[18] Hanan Farhat, George E. Sakr, Rima Kilany,“ Deep learning applications in pulmonary medical imaging: recent updates and insights on COVID-19‘‘,Machine Vision and Applications (2020).

[19] Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi,“ Review of deep learning: concepts, CNN architectures, challenges, applications, future directions“ Alzubaidi et al. J Big Data (2021) 8:53.

[20] Mohammad Mustafa Taye,“ Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions“ Computation 2023, 11, 52.

[21] Jaspreet Kaur, Prabhpreet Kaur,“ UNIConv: An enhanced U-Net based InceptionV3 convolutional model for DR semantic segmentation in retinal fundus images“, Concurrency Computat Pract Exper. 2022;e7138.

[22] Leiyu Chen, Shaobo Li, Qiang Bai,“ Review of Image Classification Algorithms Based on Convolutional Neural Networks“, Remote Sens. 2021, 13, 4712.