International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(22s) 2026 pp. 40-58
https://doi.org/10.70917/ijcisim-2026-5416
Article

An Efficient Image Segmentation Approach Using Fully Convolutional Neural Networks with Attention Fusion for Ophthalmic Disease Classification

Mr Sachin Malviya[1*], Dr. Prajeet Sharma[2]

[1*]PhD Scholar, Department of Computer science and engineering, LNCT University Bhopal, India.
Email: olsachin8@gmail.com
[2]Associate Professor, Department of Computer Science and engineering, LNCT University Bhopal, India.
Email: prajeetsharma09@gmail.com, ORCID: https://orcid.org/0009-0005-9264-7366.

Abstract: The process of segmenting retinal images and classifying ophthalmic diseases establishes a crucial pathway for both early disease identification and clinical treatment choices. Deep Learning (DL) methods demonstrate effective results however, previous research suffers from three critical issues, which include their failure to extract sufficient features, their inability to generalize well and their weak connection between preprocessing and augmentation methods. This research develops an Attention-Based Fully Convolutional Neural Network (FCNN) framework to solve existing problems in ophthalmic disease classification through Ocular Image Analysis – Ophthalmic Disease Intelligent Recognition (OIA-ODIR) retinal fundus image dataset analysis. The proposed methodology applies image preprocessing methods, which include resizing, denoising and histogram equalization, together with data augmentation techniques, which include rotation, scaling, normalization and cropping to generate new dataset samples and enhance the quality of images. The attention-based FCNN architecture classifies processed retinal images through its ability to extract retinal features that distinguish different eye conditions. The experimental results show that the developed model reaches an accuracy of 99.20% with a precision of 98.17% and a recall of 99.35%, and an F1-score value of 98.74%, which surpasses multiple current algorithms. The study concludes that the integration of preprocessing, augmentation, and attention mechanisms significantly enhances segmentation and classification performance. The proposed framework delivers accurate results that demonstrate strong performance and dependable outcomes for classifying ophthalmic diseases through retinal fundus image analysis.
Keywords: Attention-Based FCNN, Deep Learning, Ophthalmic Disease Classification, Retinal Fundus Images, Retinal Image Segmentation

1. Introduction

Image segmentation has become a fundamental requirement for ophthalmic medical image analysis since doctors need precise retinal structure depiction to make accurate disease diagnoses [1,2]. The segmentation techniques enable precise tracking of particular regions, which contain both lesions and vital anatomical components that physicians need to detect ophthalmic diseases [3–5]. DL technology has progressed rapidly, which enables FCNNs to learn spatial hierarchies while performing high-accuracy pixel-based predictions [6–9]. Attention mechanisms in these models help to identify critical features, which results in improved segmentation performance [10–12]. The field has reached a new level of progress, which allows automated retinal image analysis to deliver enhanced disease detection capabilities that protect vision health [13,14]. The DL models in Figure 1 demonstrate successful retinal vessel segmentation because their predicted outputs match the ground truth annotations, which enables accurate pixel extraction of detailed vascular structures.

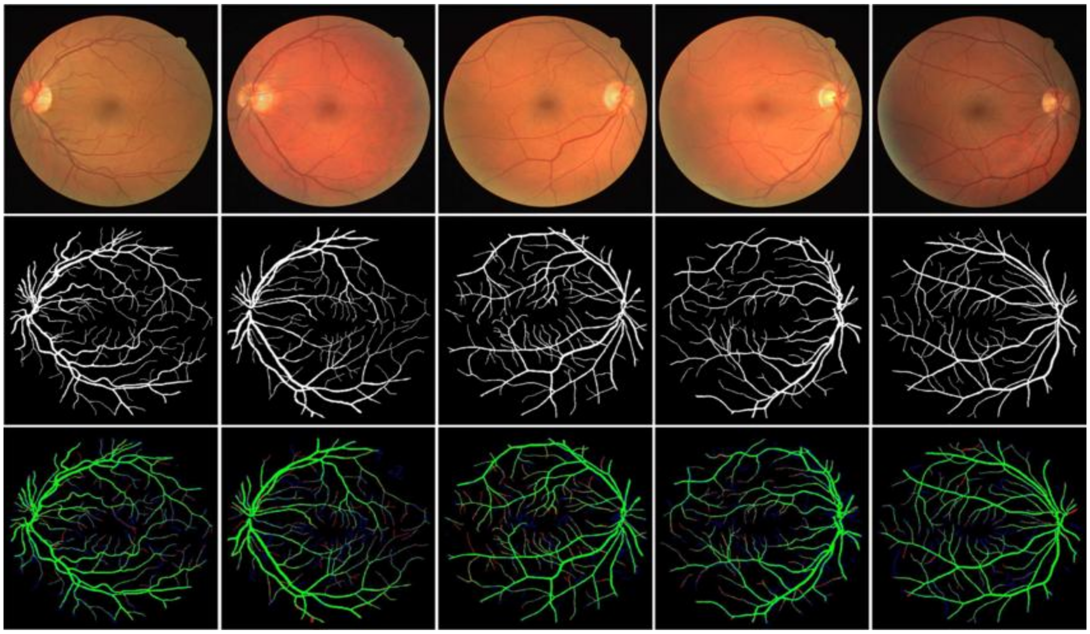


Figure 1: Retinal Image Segmentation [15].

The current research studies face multiple obstacles that prevent them from achieving accurate results in ophthalmic disease classification through segmentation methods [16-18]. The various approaches fail to handle retinal dataset differences, which arise from changes in image quality and the presence of noise and different lighting conditions [19-21]. FCNNs improved segmentation results, but these networks still have difficulty detecting small structural features and use multiscale context information poorly [22-25]. The combination of inadequate data preprocessing and weak data augmentation techniques creates conditions that lead to overfitting and decrease model accuracy on unseen data [26-29]. Researchers need to develop an all-inclusive system that combines data preprocessing with augmentation processes and advanced methods for feature extraction to solve existing difficulties [30,31].

The research introduces a novel framework that addresses current issues by implementing structured data preprocessing together with data augmentation methods and an attention-based FCNN model, which executes image segmentation and classification functions. The OIA-ODIR dataset serves as the initial dataset for the project, which then follows with image preprocessing operations that include image resizing, denoising and enhancement processes to achieve better image quality results. The dataset receives an increase in variety through data augmentation methods, which include rotation, normalization, cropping and scaling techniques. The processed data is then divided into three sets, which are used for training, validation and testing purposes. The Attention U-shaped Convolutional Network for Biomedical Image Segmentation (U-Net) architecture uses encoder and decoder paths, which include attention gates and skip connections to enable the model to identify essential areas while maintaining vital spatial information for precise segmentation and classification results.

The novelty of this research is the proven usage of attentive feature fusion over a full end-to-end pipeline, which encompasses preprocessing, augmentation and segmentation, for eye disease classification. Instead of the traditional concept, which only concentrates on designing an efficient model, this research exhibits the advantage of a superficial detailing of the data, appropriate augmentation and attention gates, in conjunction with the model for an accurate segment and classify performance. Whenever introduced to skip connections, additional gates spotlight the foreground localizations and haze the background information for a more accurate segment.

The primary goal of this research is to establish an effective, reliable segmentation-driven diagnostic scheme of ophthalmic disease classification with an attention-based fully convolution neural network. The proposed approach aims to enhance image quality and improve feature representation while achieving better classification results through its successful combination of preprocessing methods with augmentation techniques and attention mechanisms. The key contributions of this study are as follows:

- Proposed an end-to-end framework that includes preprocessing, augmentation, and an attention-based FCNN for analysis of ophthalmic images.
- Created an Attention U-Net structure to boost feature extraction and segmentation of the areas of interest in disease detection.
- Employed efficient preprocessing techniques and augmentation methods to increase model robustness and generalization.
- Shown better results in segmentation and classification than previous solutions.

The rest of the paper is structured as follows: Section II discusses the related literature, Section III explains the methodology used, Section IV provides the results and discussions, and Section V draws the conclusion and future scope.

2. Literature Review

Recent studies have extensively explored DL techniques for improving eye disease classification and retinal image analysis. **Sah et al. (2026) [32]** developed a DL framework with dual backbones, which integrated Efficient Neural Network (EfficientNet)-B0 and Residual Network (ResNet)-50 and InceptionV3 and AlexNet through multiple fusion techniques to identify different eye diseases. The system achieved 97.99% accuracy and 0.999 area under the receiver operating characteristic curve through its two fusion methods, which enhanced representation learning and generalization performance on external datasets. Similarly, **Faisal Ahmed (2026) [33]** developed a hybrid Histogram of Oriented Gradients (HOG)-Convolutional Neural Network (CNN) system, which combined manually constructed HOG features with deep CNN image analysis to identify retinal diseases, resulting in 98.5% accuracy and 99.2% Area Under the Curve (AUC) performance. **Venkataiah et al. (2025) [34]** introduced a DL system that combined Trans-MobileUNet segmentation with an attention-based dilated hybrid network to detect glaucoma, achieving 94% overall accuracy, while **Zhu et al. (2025) [35]** developed an improved HarDNet-based fully convolutional network, which employed multiscale feature extraction and dense aggregation for successful performance on CHASE_DB1 and DRIVE datasets. **Vidivelli et al. (2025) [36]** presented a DL framework based on transfer learning, which used a CNN to perform multiclass classification of eye diseases through the ODIR dataset. The system achieved its highest accuracy of 89.64% through the MobileNet model.

In addition, **Karn and Abdulla (2024) [37]** developed a hybrid attention-based U-Net model for automated segmentation of sub-retinal layers in Optical Coherence Tomography (OCT) images, incorporating edge and spatial attention mechanisms to improve feature extraction and segmentation accuracy, achieving a Dice score of 94.99% and Adjusted Rand Index (ARI) of 97.00%. **Inan (2024) [38]** introduced a hybrid Trio-model combining transfer learning CNN, two-stage CNN, and Siamese networks for multi-disease diagnosis using fundus images, achieving 97% accuracy and 0.96 AUC. Similarly, **Abdullah et al. (2024) [39]** developed a hybrid deep feature-based classification system using EfficientNet-B6 and DenseNet169 with a Decision Tree classifier, achieving 99.02% accuracy. Moreover, **Al-Fahdawi et al. (2024) [40]** proposed a multi-label DL classification system using feature-level data fusion from paired fundus images, achieving an F1-score of 88.56% and an AUC of 99.76%.

Earlier studies also contributed significantly to classification and segmentation tasks. **Mahmood et al. (2023) [41]** developed a DL model based on the Visual Geometry Group-19 architecture using transfer learning to detect cataracts through balanced binary classification on fundus images, achieving an accuracy range of 94%–95% with F1-scores between 0.95 and 0.96. Similarly, **Fernandes (2023) [42]** developed a DL framework comparing CNN and Vision Transformer models for diabetic retinopathy severity classification, where the transformer model achieved a maximum F1-score of 90 and outperformed the CNN model. Moreover, **Shamsan et al. (2023) [43]** proposed a hybrid feature-based framework integrating MobileNet and DenseNet121 with handcrafted features for multi-eye disease classification, achieving 98.5% accuracy and 99.23% AUC. In addition, **He et al. (2023) [44]** proposed a multiscale DL model based on ConvNeXt for retinal layer segmentation, achieving a Dice score of 91.3% and a mean intersection over union of 84.4%. Furthermore, **Rivas Vázquez et al. (2023) [45]** proposed a DL-based segmentation method using transfer learning with a U-Net architecture, achieving a Dice score of 83.6%. Finally, **Li et al. (2022) [46]** developed a multi-label fundus image classification model using attention mechanisms and feature fusion, achieving 94.23% accuracy and 99.16% F1-score.

2.1 Research Gaps

Although remarkable progress in DL-based analyses of ophthalmological data has been achieved, there still remain some open questions in this area. First of all, a lack of investigation regarding hybrid schemes that utilize both image segmentation and classification or feature extraction along with their fusion in the same framework is observed in [32,34,44]. Secondly, there is no adequate study about robust methods of feature fusion, which may allow effective processing of multimodal and multiscale information about retinal diseases [33,43]. In addition, generalization of approaches in order to ensure the ability of models to achieve good results regardless of the dataset used is still insufficient [35,36]. Moreover, investigation of simultaneous multiclass classification/segmentation is still needed [39,46].

3. Research Methodology

This section describes the methodology that is presented in Figure 2. It involves the use of the Attention-Based FCNN technology in order to develop an efficient image segmentation system. The whole process of the research consists of five major actions. These include data set preparation, data preprocessing, data augmentation, model development, and performance evaluation. The suggested approach is based on the use of the attention-based FCNN network.

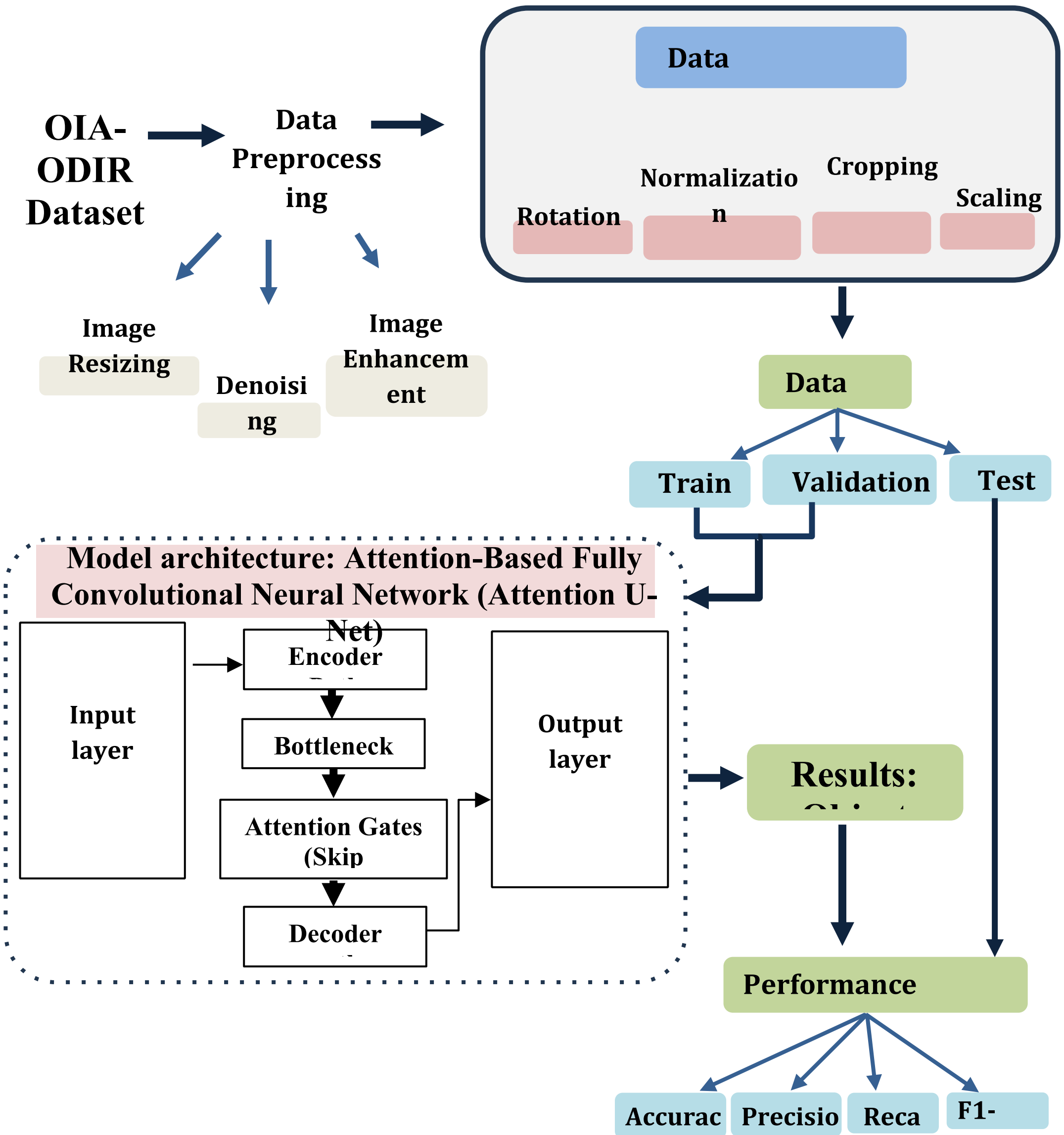


Figure 2: Proposed Methodology

3.1 Dataset Description

The proposed research utilizes a secondary dataset acquired from the Kaggle platform, consisting of the OIA-ODIR retinal fundus image dataset [47]. The dataset contains retinal fundus images that were collected from approximately 5000 patients who provided images of both their left and right eyes, which resulted in a total of about 10000 images. The dataset contains various real-world variations, which include different lighting conditions, image quality and disease states, which makes it useful for testing segmentation models in difficult situations. Table 1 delivers a comprehensive overview of all dataset attributes.

Table 1: Dataset description

Parameter	Description
Dataset Name	OIA-ODIR
Publisher / Source	Microsoft Research

Dataset Type	Secondary
Total Images	10,000 retinal fundus images
Disease Categories	Eight ophthalmic disease categories, including normal eyes and various ocular diseases
Annotation Types	Image-level disease labels (no ground-truth pixel-level masks provided)
Image Resolution	Varies across images (commonly ranging from approximately 480×640 to higher resolutions)
Format	Joint Photographic Experts Group (JPEG) images
Advantages	Provides clinically diverse real-world retinal images, enabling robust learning of disease-specific features and improving generalization performance

3.2 Image Preprocessing

Once the dataset is acquired, the input images are preprocessed [48] to enhance their quality and normalize them. Let the input image I. The first manipulation is to resize the image to a fixed spatial resolution so that the inputs are of uniform size, which is written as:

$$I_r = R(I) \quad (1)$$

where $R(\cdot)$ denotes the resizing function and I_r is the resized image.

A denoising operation is utilized first to reduce artifacts and noise in the input images using a smoothing filter. This process can be defined as:

$$I_d = I_r * G_\sigma \quad (2)$$

where G_σ represents the Gaussian kernel and $*$ denotes the convolution operation, resulting in the denoised image I_d.

The process of image enhancement begins with the goal of increasing contrast to show essential details of the image. The following equation describes the transformation:

$$I_e = E(I_d) \quad (3)$$

where $E(\cdot)$ represents the enhancement function and I_e is the enhanced output image.

The preprocessing operations work together to enhance input images because they diminish noise and boost visual quality while creating uniform image standards that enable successful feature extraction.

3.3 Data Augmentation

Data augmentation [49] follows preprocessing because it increases input data variation while building model strength. The preprocessed image exists as I_e. The first step involves normalization, which scales the pixel intensity values to a standard range and is defined as:

$$I_n = \frac{I_e - I_{min}}{I_{max} - I_{min}} \quad (4)$$

where I_{min} and I_{max} denote the minimum and maximum pixel intensity values, respectively, and I_n is the normalized image.

To introduce variability in spatial orientation, rotation is applied as:

$$I_{rot} = T_\theta(I_n) \quad (5)$$

where I_{rot} represents the rotated image, $T_\theta(\cdot)$ denotes the rotation transformation, I_n is the normalized input image, and θrepresents the rotation angle.

Further, scaling is performed to simulate variations in object size:

$$I_s = S_\alpha(I_{rot}) \tag{6}$$

where I_s represents the scaled image, $S_\alpha(\cdot)$ denotes the scaling transformation, I_{rot} is the rotated image, and $\alpha$is the scaling factor.

Finally, cropping is applied to extract relevant regions from the transformed image:

$$I_c = C(I_s) \tag{7}$$

where I_c represents the cropped image, $C(\cdot)$ denotes the cropping operation, and I_s is the scaled image.

The augmentation operations create different versions of the input data, which enable the model to acquire two types of features, invariant features and discriminative features, while enhancing its performance across various orientations, scales and spatial distribution changes.

3.4 Data Splitting

Following data augmentation, the dataset is subjected to three types of data splitting processes, which help in the creation and validation of the model. Generally, the dataset is usually divided into three categories: training, validation, and testing data in the proportion of 70:15:15, respectively. The training data facilitates model parameter learning, while the validation data helps in monitoring the training process and adjusting the hyperparameters accordingly. Lastly, the testing data facilitates the assessment of the model, which helps gauge the prediction capability of the model using new data.

3.5 Model Architecture – Attention-Based FCNN

The proposed model uses an Attention-Based FCNN [50], which functions as a pixel-level image segmentation system. The basic architecture of the FCNN is shown in Figure 3, which consists of an input layer, multiple convolutional (hidden) layers for feature extraction, and an output layer that produces pixel-wise predictions. The model uses its layered design to develop multiple levels of image understanding, which help it to achieve precise segmentation results.

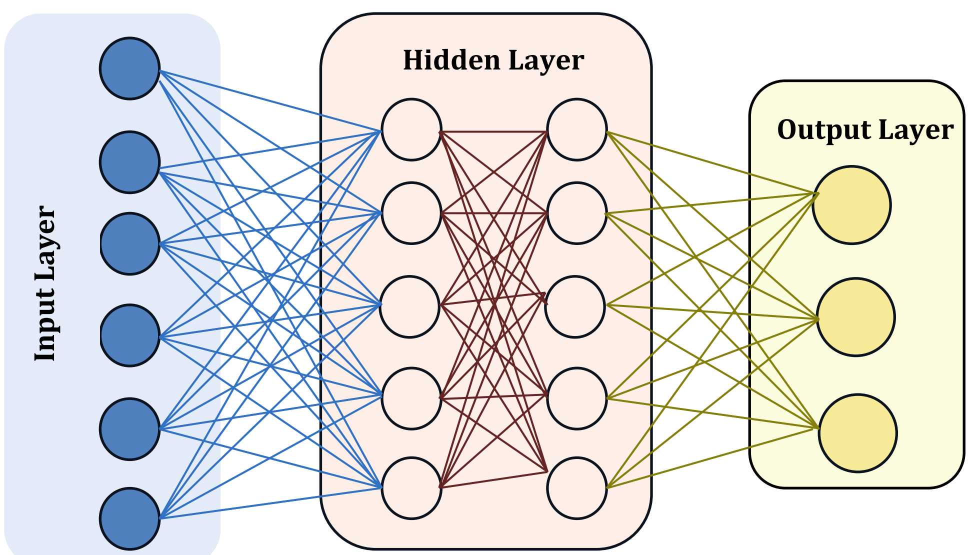


Figure 3: FCNN architecture [51]

In the convolutional layers, features are extracted in a hierarchical way from the given input image, and every single layer would learn a spatial and contextual representation of the image. Deepening of the network produces more semantic features while lowering the spatial resolution.

The system uses intermediate feature representations to extract essential semantic information, which is used to create accurate segmentation results. The network keeps all spatial information through feature propagation because it needs to preserve both its basic and complex spatial details.

The proposed model uses attention gates, which enable the network to focus on important areas while it discards unnecessary background content. The attention mechanism is defined by the following mathematical expression:

$$\alpha = \sigma(W_x x_l + W_g g + b) \quad (8)$$

The attention coefficients are represented by α while x_l denotes the feature map corresponding to layer l, and g functions as the gating signal and W_x and W_g function as learnable weight parameters, and b acts as the bias termandσserves as the activation function. The refined feature map is obtained as:

$$\hat{x}_l = \alpha \cdot x_l \quad (9)$$

The attention-weighted feature map is represented by $\hat{x}_l$. The model uses this mechanism to identify crucial spatial areas, which results in better segmentation performance.

Finally, the model creates pixel-wise class probabilities through the SoftMax function, which enables pixel classification into specific classes to generate the final segmented result.

3.6 Results – Object Classification

The system produces pixel-level predictions, which serve for both object classification and segmentation after it processes through the Attention-Based FCNN model. The trained model produces a segmentation map in which each pixel is assigned to a specific class, representing different regions of interest within the image. The output provides a structured representation by clearly distinguishing between objects and background at the pixel level, enabling accurate localization of important features. The proposed method successfully captures all spatial and contextual details, which enable precise image segmentation and classification.

3.7 Performance Evaluation

The evaluation of the proposed Attention-Based FCNN model performance employs conventional metrics quantifying its ability to correctly predict class labels by comparing the predicted values with true, ground truth data of the ground experiments. The true positives, true negatives, false positives and false negatives are denoted by TP, TN, FP and FN, respectively. The metrics are given by:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10) \qquad Precision = \frac{TP}{TP+FP}$$

$$(11) \quad Recall = \frac{TP}{TP+FN} \qquad (12) \quad F1 - score = \frac{2\cdot(Precision\cdot Recall)}{Precision+Recall}$$

(13)

3.8 Proposed Algorithm

Algorithm: Effective Image Segmentation using FCNN

Input: OIA-ODIR Dataset with annotated images

Output: Segmented images with pixel-level class labels

BEGIN

Load OIA-ODIR dataset

```
FOR every image in the dataset
    Resize image to uniform characteristic (e.g., 256x256)
    Apply a denoising filter (Gaussian or Median)
    Perform image enhancement (Histogram Equalization)
END FOR
// Data Augmentation
FOR each image in the dataset DO
   Normalize pixel values between [0,1]
   Apply random rotation, scaling, and cropping
   Store augmented images in the training set
END FOR
Split dataset → Training (70%), Validation (15%), Testing (15%)
// Model Definition
Apply attention gate to encoder feature maps using decoder gating signal before feature concatenation.
Initialize the Attention-Based FCNN model.
Define Encoder path: Convolution + ReLU + Max Pooling layers.
Define Bottleneck layer: High-level feature extraction
Define Attention Gates: Skip connections
Define Decoder path: Upsampling + Convolution
Add Output layer: SoftMax activation for pixel-wise classification
// Model Training
Compile model with optimizer (Adam) and loss function (Cross-Entropy)
Train Attention-Based FCNN on training data with batch size B for N epochs
Validate the model on the validation set after each epoch
Keep the best-performing model bulks
// Model Testing and Evaluation
Load test data and perform a prediction
Compare predicted segmentation maps with ground truth
Compute performance metrics
Display segmented outputs and performance graphs
END
```

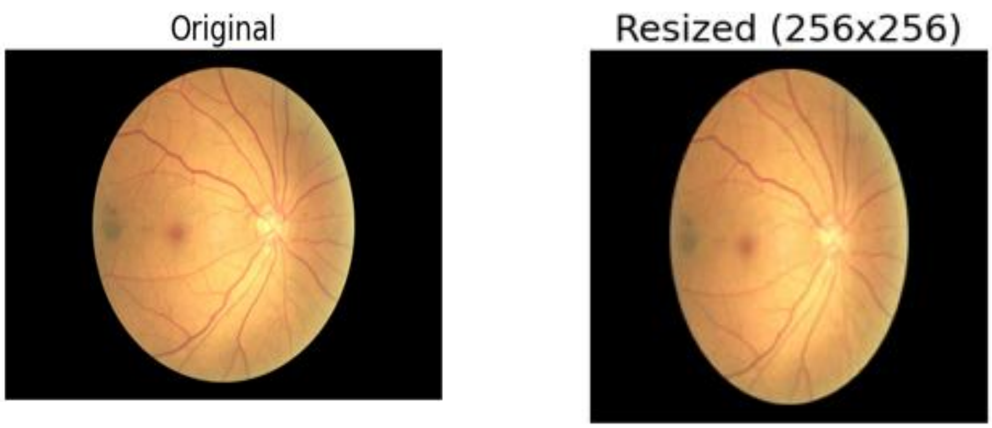
4. Results and Discussion

The following part of the paper contains experimental results and performance analysis of the Attention-Based FCNN model used for the segmentation and classification of ocular diseases from images. The proposed framework proves itself effective by performing performance analysis that takes into account the values of accuracy, precision, recall, and F1-Score. In Table 2, the hyperparameters required for the training of the model are provided.

Table 2: Hyperparameters Used

Hyperparameter	Value
Hidden layer dimension	512
Fundus projection	Linear (fundus_dim → 512)
OCT projection	Linear (oct_dim → 512)
Attention MLP layers	2
Attention output size	2
Activation function	ReLU
Attention normalization	Softmax
Output dimension	512

The OIA-ODIR retinal fundus images, which have been obtained, undergo preprocessing procedures that improve their visual quality while maintaining their required input size dimensions, as shown in Figure 4. The original retinal images undergo resizing to 256×256 resolution, which establishes uniformity throughout the dataset. Denoising procedures follow this step to eliminate all undesired elements, which include background noise and other artifacts found in the fundus images. The process uses histogram equalization to improve image brightness, which enables better detection of retinal vessels and important eye structures, thus resulting in more informative retinal images.



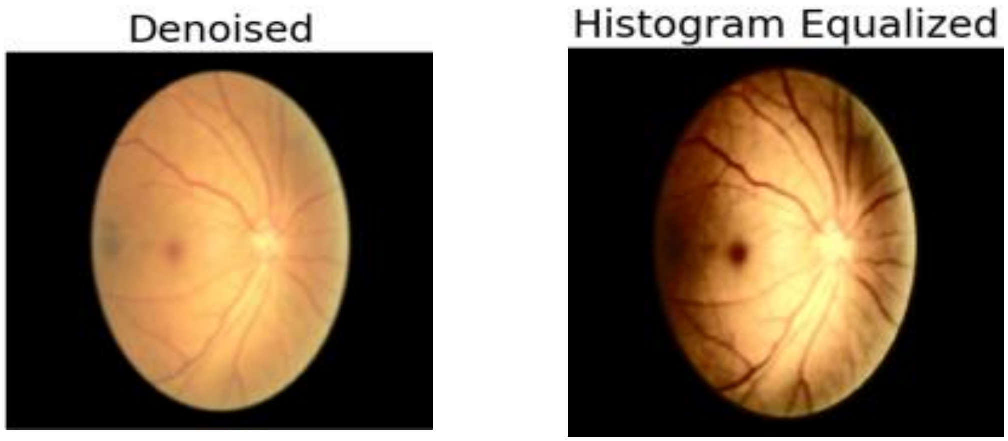


Figure 4: Image Preprocessing Stages.

The retinal fundus images, which have been preprocessed, need additional enhancements through data augmentation to generate diverse data that would strengthen the performance of machine learning models. Figure 5 shows various augmented samples that have been generated from the original image by applying transformations such as rotation, scaling, and spatial variations.

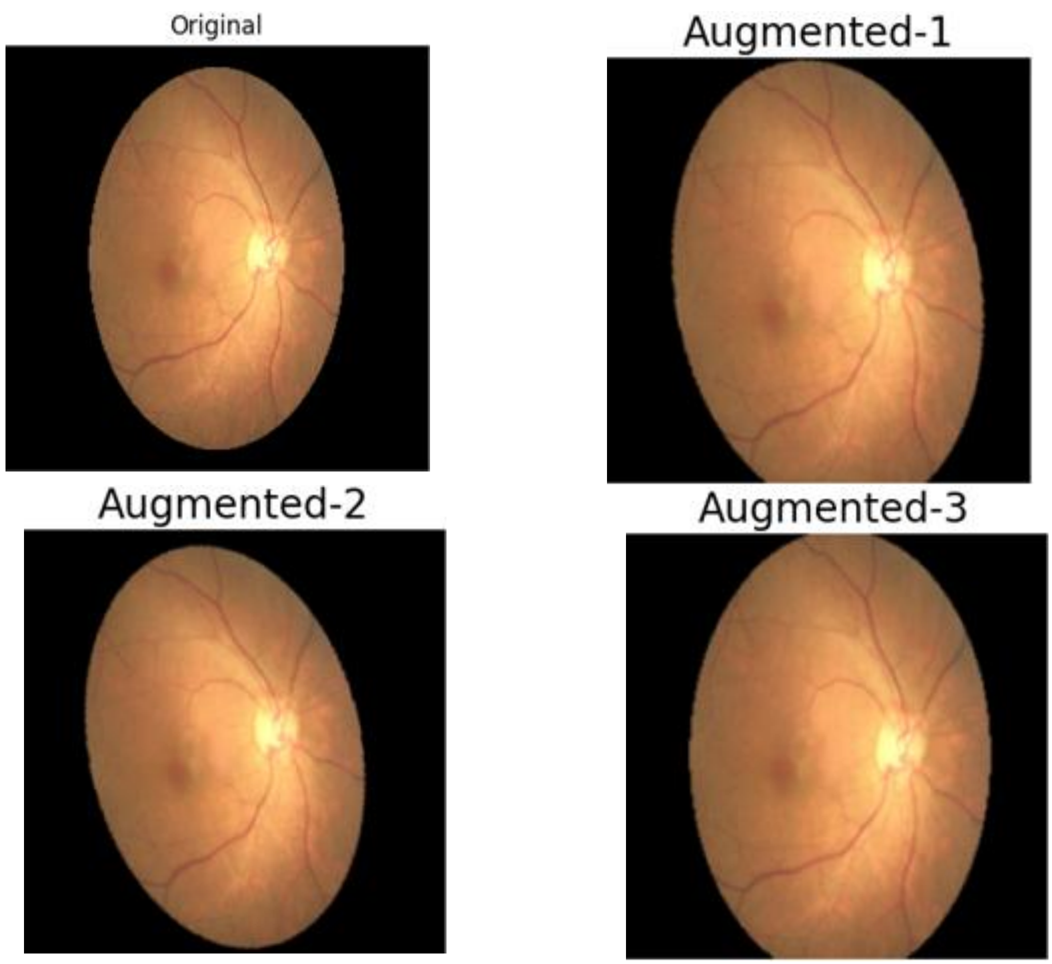


Figure 5: Data Augmentation Results for Retinal Image

The dataset is divided into training, validation, and testing sets to ensure balanced model learning and performance evaluation. The dataset distribution in Table 3 shows that 70% of the images, which equals 4900 images, are used for training, while 15% of the images, which equals 1050 images, are used for both validation and testing.

Table 3: Dataset Partitioning for Model Development and Evaluation

Dataset Category	**Percentage (%)**	**Number of Images**
Training Set	70%	4900
Validation Set	15%	1050
Testing Set	15%	1050
Total	100%	7000

The proposed Attention-Based FCNN model demonstrates its training loss performance through Figure 6, which shows the results from five training epochs. The loss value decreases from 1.51 in the first epoch to 0.31 in the last epoch, which shows the model achieved stable learning progress during training. The model demonstrates effective retinal feature learning through its gradual loss reduction, which shows its ability to classify and segment with minimal errors.

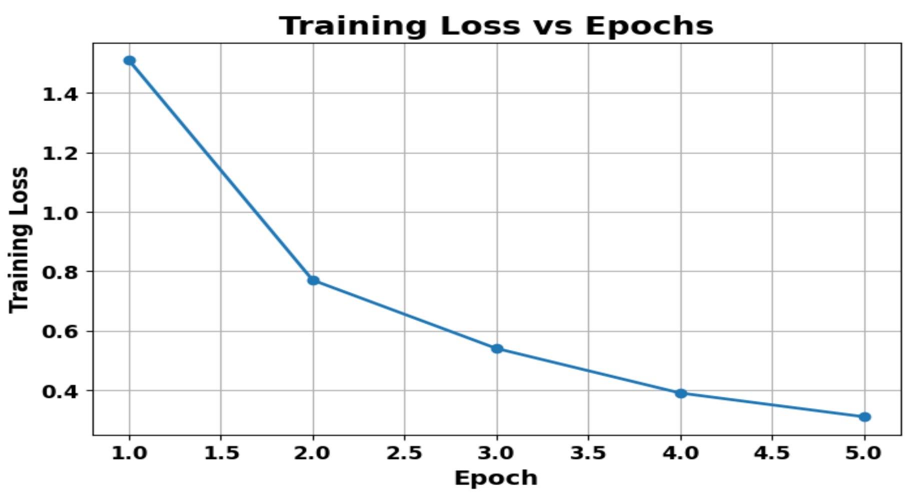


Figure 6: Training Loss vs Epochs

Figure 7 illustrates the validation accuracy of the Attention-Based FCNN model proposed, which has been trained for five epochs. The figure shows that the validation accuracy regularly increased from 85.6% at epoch 1 to 98.9% at epoch 5. The gradually increased validation accuracy suggests the good generalization of retinal images using the proposed model with a stable classification and segmentation capability.

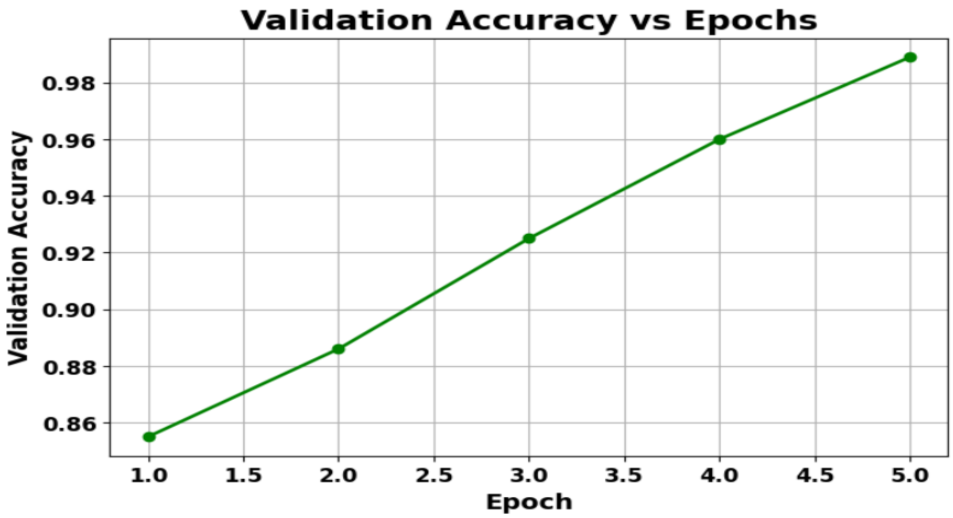


Figure 7: Validation Accuracy vs Epochs

Figure 8 shows the classification result of the proposed Attention-based FCNN for a sample retinal fundus image is provided by the proposed classifier. The proposed classifier can identify class C with the highest likelihood (40.51%) and other classes A and O with the likelihood scores of 24.10% and 17.29%. These probability scores indicate the proposed model can certainly discriminate ophthalmic condition spaces using learned retinal images.

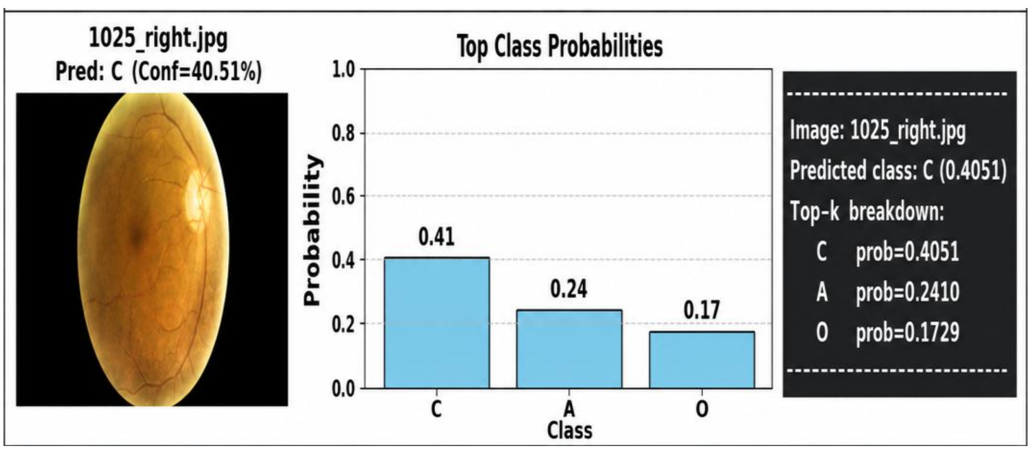


Figure 8: Class Probability Distribution for Sample Prediction

Figure 9 illustrates another classification result of the proposed Attention-Based FCNN model applied to a retinal fundus image. The model successfully identified the image as class C with 38.43% highest class probability, together with other probabilities of class A and O as 22.66% and 17.82%, respectively. It is observed that the probability of each class is reasonable for grading classification, suggesting consistent classification as well as retinal disease feature extraction from the input image.

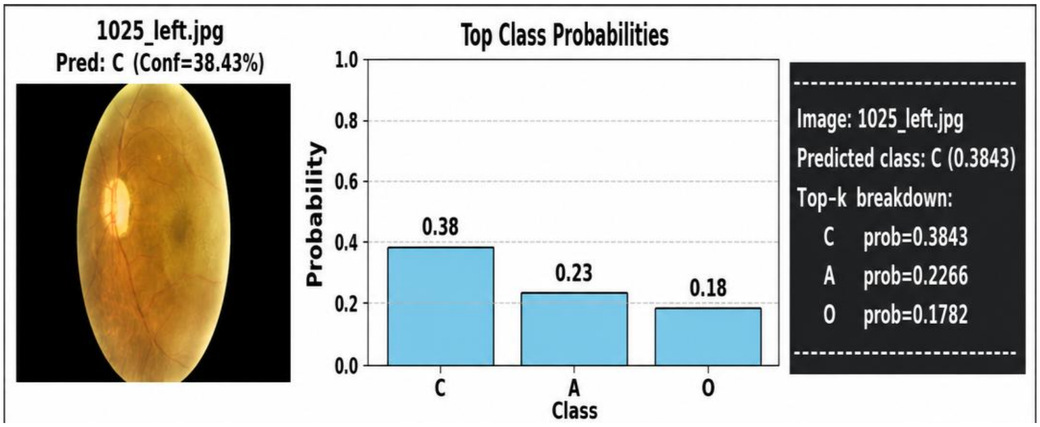


Figure 9: Class Probability Distribution for Sample Image

The confusion matrix of the developed attention-based FCNN classifier has been illustrated in Figure 10 based on test data with 99.20% classification accuracy. It can be seen from the confusion matrix that most of the samples are correctly classified with respect to their diagonals, such as 647 samples belong to category D, 644 samples belong to category N, and 308 samples belong to category O. From the results, it is clear that the proposed approach has robustness in disease classification in ophthalmology.

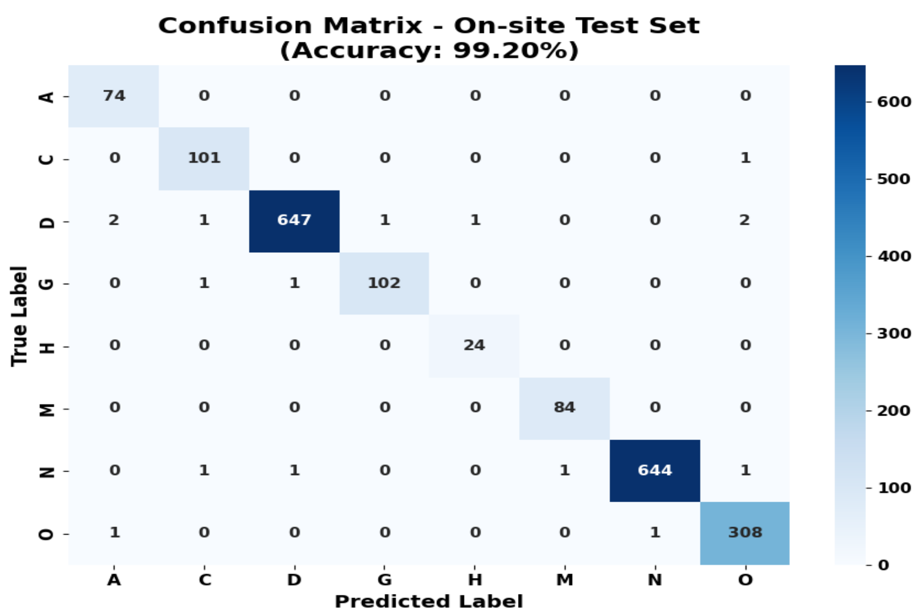


Figure 10: Confusion Matrix

The proposed Attention-Fusion FCNN model shows its class-specific performance results through the execution metrics in Table 4. The model demonstrates consistently high precision, recall, and F1-score values across all eight classes, with most metrics ranging between 98% and 100%. The precision values for Classes D and N reach 100% while Classes A, H, and M achieved 100% recall, which proves their capacity to identify real instances. The F1-scores maintain a constant range of 98 to 100 percent, which indicates that the system achieves balanced classification results through its minimal false positive and false negative errors. The proposed multimodal attention-based framework achieves 99.20 percent overall accuracy, which establishes its dependable performance across all classification categories.

Table 4: Class-wise performance of the proposed model

Class	Precision	Recall	F1-Score	Accuracy
A	96%	100%	98%	99.20%
C	97%	99%	98%	
D	100%	99%	99%	
G	99%	98%	99%	
H	96%	100%	98%	
M	99%	100%	99%	
N	100%	99%	100%	
O	99%	99%	99%	

The performance results show the operational efficiency of the Attention-Based FCNN model through the measurement results shown in Figure 11. The model achieves an accuracy of 99.20%, precision of 98.17%, recall of 99.35%, and F1-score of 98.74%, which demonstrates its ability to perform balanced assessment across different categories of ophthalmic diseases.

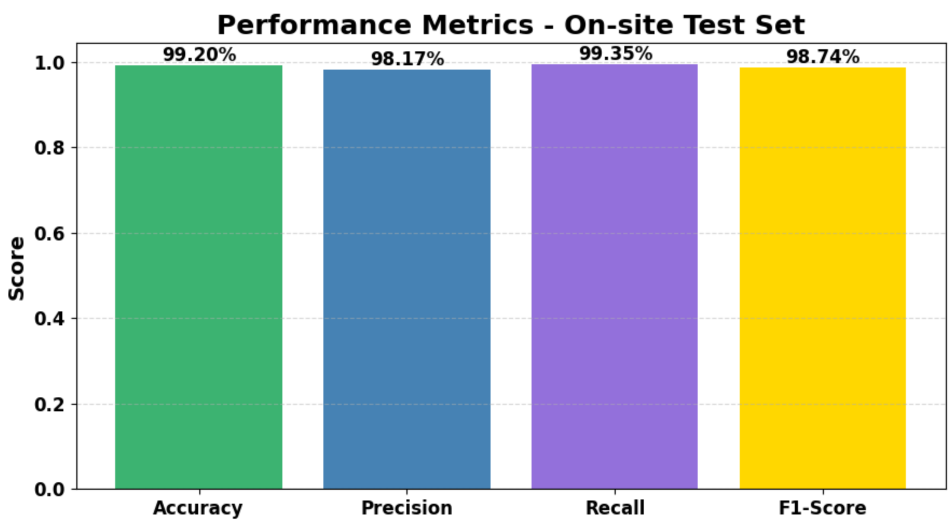


Figure 11: Overall Performance Metrics

4.1 Ablation Study

The ablation study results of the proposed Attention-Based FCNN model show different model configurations in Table 5. The results show that the integration of preprocessing and data augmentation gradually improves the overall classification performance, where the proposed model achieves the highest accuracy of 99.20%, precision of 98.17%, recall of 99.35%, and F1-score of 98.74%, demonstrating the effectiveness of the complete framework for ophthalmic disease classification.

Table 5: Ablation Study

Model Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
FCNN	91.42	90.85	91.10	90.96
FCNN with preprocessing	94.78	94.11	94.52	94.30

FCNN with preprocessing and augmentation	96.95	96.37	96.84	96.58
Proposed Attention-Based FCNN Model	99.20	98.17	99.35	98.74

4.2 Comparative Analysis

Table 6 presents the comparative analysis of the proposed Attention-Based FCNN model with existing ophthalmic disease classification approaches. Sah et al. (2026) [32] combined EfficientNet-B0 with ResNet50, InceptionV3, and AlexNet to achieve 97.99% accuracy, while Faisal Ahmed (2026) [33] proposed the HOG-CNN model with 98.5% accuracy. Sakib Rokoni (2025) [52] achieved 87.37% accuracy using the Hybrid LsEyeNet model, whereas Al-Fehdawi et al. (2024) [40] obtained 89.18% accuracy with Fundus-DeepNet. Muhammed Pektas (2023) [53] utilized EfficientNetB3 and achieved 96.64% accuracy. Compared with these existing approaches, the proposed Attention-Based FCNN model achieves the highest classification accuracy of 99.20%, demonstrating improved ophthalmic disease classification performance.

Table 6: Comparative Analysis

Author	Technique Used	Accuracy (%)
Sah et al. (2026) [32]	EfficientNet-B0 with ResNet50, InceptionV3, and AlexNet	97.99
Faisal Ahmed (2026) [33]	HOG-CNN	98.5
Sakib Rokoni (2025) [52]	Hybrid LsEyeNet model	87.37
AI-Fehdawi et al. (2024) [40]	Fundus-DeepNet	89.18
Muhammed Pektas (2023) [53]	EfficientNetB3	96.64
Proposed study (2026)	Attention-Based FCNN	99.20

4.3 Discussion

The experimental results show that the Attention-Based FCNN model achieves better classification results for ophthalmic diseases through its combination of preprocessing methods, data augmentation techniques, and attention-based feature learning methods. The preprocessing stage produces better retinal image results because it reduces noise and improves visibility through contrast enhancement. The training and validation results demonstrate stable performance through decreasing loss values and rising accuracy levels throughout the training process. The classification outputs together with the confusion matrix demonstrate that the proposed model has the ability to correctly identify various ophthalmic disease categories while making only a few errors. The model achieves 99.20% accuracy, 98.17% precision, 99.35% recall, and 98.74% F1-score, which demonstrates its ability to classify data with high accuracy and the same performance level across different categories. The ablation study also demonstrates that the combination of preprocessing and augmentation techniques delivers essential benefits for improving system performance. The comparative analysis shows that the proposed Attention-Based FCNN model exceeds the performance of multiple current DL methods used for ophthalmic disease classification.

5. Conclusion and Future Scope

Retinal image analysis serves as a crucial method for detecting ophthalmic diseases at their initial stages because it requires precise retinal structure segmentation and classification for proper clinical evaluations. The existing DL methods from earlier studies have numerous shortcomings because they fail to extract important features and to generalize well, while their combination of preprocessing methods and data augmentation techniques remains inadequate. The study developed an Attention U-Net system to solve these problems because they needed an effective method for classifying medical conditions that used the OIA-ODIR retinal fundus image dataset. The proposed methodology uses a sequence that begins with image preprocessing and data augmentation, followed by dataset splitting and concludes with attention-based feature learning, which produces better results for both segmentation and classification tasks. The experimental test results show that the proposed model reaches stable convergence to achieve classification results with 99.20% accuracy, and 98.17% precision, and 99.35% recall and 98.74% F1-score. Comparative analysis shows that the proposed method achieves better performance than multiple existing DL systems that classify ophthalmic diseases. In future work, the proposed framework can be extended using larger multimodal retinal datasets and advanced transformer-based attention mechanisms to further improve classification robustness and real-world clinical applicability.

REFERENCES

1. Lai, W. E. I., and H. U. Menghan. "A review of medical ocular image segmentation." *Virtual Reality & Intelligent Hardware* 6, no. 3 (2024): 181-202. https://doi.org/10.1016/j.vrih.2024.04.001
2. Alexopoulos, Palaiologos, Chisom Madu, Gadi Wollstein, and Joel S. Schuman. "The development and clinical application of innovative optical ophthalmic imaging techniques." *Frontiers in medicine* 9 (2022): 891369. https://doi.org/10.3389/fmed.2022.891369
3. Quintana-Quintana, Oliver Jonathan, Marco Antonio Aceves-Fernández, Jesús Carlos Pedraza-Ortega, Gendry Alfonso-Francia, and Saul Tovar-Arriaga. "Deep Learning Techniques for Retinal Layer Segmentation to Aid Ocular Disease Diagnosis: A Review." *Computers* 14, no. 8 (2025): 298. https://doi.org/10.3390/computers14080298
4. Buric, Matija, Sinisa Grozdanic, and Marina Ivasic-Kos. "Diagnosis of ophthalmologic diseases in canines based on images using neural networks for image segmentation." *Heliyon* 10, no. 19 (2024). https://doi.org/10.1016/j.heliyon.2024.e38287
5. Raposo, Hugo. "Intelligent imaging: A systematic review of artificial intelligence techniques in disease detection, segmentation, and classification." *Segmentation and Classification (May 13, 2024)* (2024). http://dx.doi.org/10.2139/ssrn.5342645
6. Arosio, Riccardo, Brandon Hobley, Andrew J. Wheeler, Fabio Sacchetti, Luis A. Conti, Thomas Furey, and Aaron Lim. "Fully convolutional neural networks applied to large-scale marine morphology mapping." *Frontiers in Marine Science* 10 (2023): 1228867. https://doi.org/10.3389/fmars.2023.1228867
7. Abdellaoui, Cherifa, Samia Belkacem, and Noureddine Messaoudi. "Deep learning architectures for medical image segmentation: an organized analysis of CNN-based models and uses." *Bulletin of Electrical Engineering and Informatics* 15, no. 1 (2026): 424-437. https://doi.org/10.11591/eei.v15i1.10839
8. Dumbrique, Jakov Ivan S., Reynan B. Hernandez, Juan Miguel L. Cruz, Ryan M. Pagdanganan, and Prospero C. Naval Jr. "Pneumothorax detection and segmentation from chest X-ray radiographs using a patch-based fully convolutional encoder-decoder network." *Frontiers in Radiology* 4 (2024): 1424065. https://doi.org/10.3389/fradi.2024.1424065
9. Shivnani, Honey, Prerna Agrawal, and Savita Gandhi. "Intelligent Imaging: Deep Learning in Medical Science." In *AI in Clinical Diagnosis, Prediction, and Patient Care*, pp. 55-90. IGI Global Scientific Publishing, 2026. https://doi.org/10.4018/979-8-2600-0986-4.ch002
10. Lyu, Yuchao, Yinghao Xu, Xi Jiang, Jianing Liu, Xiaoyan Zhao, and Xijun Zhu. "AMS-PAN: Breast ultrasound image segmentation model combining attention mechanism and multiscale features." *Biomedical Signal Processing and Control* 81 (2023): 104425. https://doi.org/10.1016/j.bspc.2022.104425
11. Gonçalves, Tiago, Isabel Rio-Torto, Luís F. Teixeira, and Jaime S. Cardoso. "A survey on attention mechanisms for medical applications: are we moving toward better algorithms? " *Ieee Access* 10 (2022): 98909-98935. https://doi.org/10.1109/ACCESS.2022.3206449
12. Chen, Chen, Cheng Chen, Mingrui Ma, Xiaojian Ma, Xiaoyi Lv, Xiaogang Dong, Ziwei Yan, Min Zhu, and Jiajia Chen. "Classification of multi-differentiated liver cancer pathological images based on deep learning attention mechanism." *BMC Medical Informatics and Decision Making* 22, no. 1 (2022): 176. https://doi.org/10.1186/s12911-022-01919-1
13. Soomro, Toufique Ahmed, Faisal Bin Ubaid, Shehnila Zardari, Ahmed Ali Shah, Marwah M. Alansari, and Shimaa A. Nagro. "The state of retinal image analysis: Deep learning advances and applications." *IEEE Access* 13 (2025): 65066-65093. https://doi.org/10.1109/ACCESS.2025.3556765
14. Wu, Pan, Xiaoqiang He, Wenhao Dai, Jingwei Zhou, Yutong Shang, Yourong Fan, and Tao Hu. "A Review on Research and Application of AI-based Image Analysis in the field of Computer Vision." *IEEE Access* (2025). https://doi.org/10.1109/ACCESS.2025.3565300
15. Xie, Yutong, Bing Yang, Qingbiao Guan, Jianpeng Zhang, Qi Wu, and Yong Xia. "Attention mechanisms in medical image segmentation: A survey." arXiv preprint arXiv:2305.17937 (2023). https://doi.org/10.48550/arXiv.2305.17937
16. Rajarajeshwari, G., and G. Chemmalar Selvi. "Application of artificial intelligence for classification, segmentation, early detection, early diagnosis, and grading of diabetic retinopathy from fundus retinal images: a comprehensive review." *IEEE Access* 12 (2024): 172499-172536. https://doi.org/10.1109/ACCESS.2024.3494840

17. Vadduri, Maneesha, and P. Kuppusamy. "Enhancing ocular healthcare: deep learning-based multiclass diabetic eye disease segmentation and classification." *IEEe Access* 11 (2023): 137881-137898. https://doi.org/10.1109/ACCESS.2023.3339574
18. Kv, Rajitha, Keerthana Prasad, and Prakash Peralam Yegneswaran. "Segmentation and classification approaches of clinically relevant curvilinear structures: A review." *Journal of Medical Systems* 47, no. 1 (2023): 40. https://doi.org/10.1007/s10916-023-01927-2
19. Shi, Chuying, Jack Lee, Gechun Wang, Xinyan Dou, Fei Yuan, and Benny Zee. "Assessment of image quality on color fundus retinal images using the automatic retinal image analysis," *Scientific Reports* 12, no. 1 (2022): 10455. https://doi.org/10.1038/s41598-022-13919-2
20. Celik, Canan, and İbrahim Yücedağ. "The role of deep learning-based algorithms in retinal image processing: a review." *Multimedia Tools and Applications* 85, no. 3 (2026): 208. https://doi.org/10.1007/s11042-026-21181-1
21. Li, Heng, Haojin Li, Mingyang Ou, Xiangyang Yu, Xiaoqing Zhang, Ke Niu, Huazhu Fu, and Jiang Liu. "Fundus image quality assessment and enhancement: a systematic review." *arXiv preprint arXiv:2501.11520* (2025). https://doi.org/10.48550/arXiv.2501.11520
22. Liu, Ye, Huifang Li, Chao Hu, Shuang Luo, Yan Luo, and Chang Wen Chen. "Learning to aggregate multiscale context for instance segmentation in remote sensing images." *IEEE Transactions on Neural Networks and Learning Systems* 36, no. 1 (2024): 595-609. 10.1109/TNNLS.2023.3336563
23. Zhao, Xiaoqi, Hongpeng Jia, Youwei Pang, Long Lv, Feng Tian, Lihe Zhang, Weibing Sun, and Huchuan Lu. "M ^{2} SNet: Multiscale in multiscale subtraction network for medical image segmentation." arXiv preprint arXiv:2303.10894 (2023) https://doi.org/10.1007/s11633-026-1662-9
24. Wu, Honglin, Min Zhang, Peng Huang, and Wenlong Tang. "CMLFormer: CNN and multiscale local-context transformer network for remote sensing images semantic segmentation." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 17 (2024): 7233-7241. 10.1109/JSTARS.2024.3375313
25. He, Along, Kai Wang, Tao Li, Wang Bo, Hong Kang, and Huazhu Fu. "Progressive multiscale consistent network for multiclass fundus lesion segmentation." *IEEE transactions on medical imaging* 41, no. 11 (2022): 3146-3157. 10.1109/TMI.2022.3177803
26. Ju, Zedong, Yinsheng Chen, Yukang Qiang, Xinyi Chen, Chao Ju, and Jingli Yang. "A systematic review of data augmentation methods for intelligent fault diagnosis of rotating machinery under limited data conditions." *Measurement Science and Technology* 35, no. 12 (2024): 122004. https://doi.org/10.1088/1361-6501/ad7a97
27. Wei, Sizhou, Zhiyuan Chen, Senthil Kumar Arumugasamy, and Irene Mei Leng Chew. "Data augmentation and machine learning techniques for control strategy development in bio-polymerization process." *Environmental Science and Ecotechnology* 11 (2022): 100172. https://doi.org/10.1016/j.ese.2022.100172
28. Aliferis, Constantin, and Gyorgy Simon. "Overfitting, underfitting and general model overconfidence and under-performance pitfalls and best practices in machine learning and AI." *Artificial intelligence and machine learning in health care and medical sciences: Best practices and pitfalls* (2024): 477-524. https://doi.org/10.1007/978-3-031-39355-6_10
29. Zubair Rahman, AMJ MD, R. Mythili, K. Chokkanathan, T. R. Mahesh, K. Vanitha, and Temesgen Engida Yimer. "Enhancing image-based diagnosis of gastrointestinal tract diseases through deep learning with EfficientNet and advanced data augmentation techniques." *BMC Medical Imaging* 24, no. 1 (2024): 306. https://doi.org/10.1186/s12880-024-01479-y
30. Ikegwu, Anayo Chukwu, Henry Friday Nweke, and Chioma Virginia Anikwe. "Recent trends in computational intelligence for educational big data analysis." *Iran Journal of Computer Science* 7, no. 1 (2024): 103-129. https://doi.org/10.1007/s42044-023-00158-5
31. Ikegwu, Anayo Chukwu, Henry Friday Nweke, Chioma Virginia Anikwe, Uzoma Rita Alo, and Obikwelu Raphael Okonkwo. "Big data analytics for data-driven industry: a review of data sources, tools, challenges, solutions, and research directions." Cluster Computing 25, no. 5 (2022): 3343-3387. https://doi.org/10.1007/s10586-022-03568-5
32. Sah, Uday Kumar, Jyotir Moy Chatterjee, and R. Sujatha. "Multiclass eye disease classification using deep learning EfficientNetB0 fusion techniques." Scientific Reports 16, no. 1 (2026): 6368. https://doi. org/10.1038/s41598-026-35357-0
33. Ahmed, Faisal. "Hog-cnn: Integrating histogram of oriented gradients with convolutional neural networks for retinal image classification." Journal of Imaging Informatics in Medicine (2026): 1-16. https://doi.org/10.1007/s10278-026-01893-4
34. Venkataiah, C., M. Chennakesavulu, Y. Mallikarjuna Rao, B. Janardhana Rao, G. Ramesh, J. Sofia Priya Dharshini, and Manjula Jayamma. "A novel eye disease segmentation and classification model using an advanced deep learning network." Biomedical Signal Processing and Control 105 (2025): 107565. https://doi.org/10.1016/j.bspc.2025.107565
35. Zhu, Yuanpei, Yong Liu, and Xuezhi Zhou. "Automated segmentation of retinal vessel using HarDNet fully convolutional networks." PLoS One 20, no. 9 (2025): e0330641. https://doi.org/10.1371/journal.pone.0330641
36. Vidivelli, S., P. Padmakumari, C. Parthiban, A. DharunBalaji, R. Manikandan, and Amir H. Gandomi. "Optimising deep learning models for ophthalmological disorder classification." Scientific Reports 15, no. 1 (2025): 3115. https://doi.org/10.1038/s41598-024-75867-3
37. Karn, Prakash Kumar, and Waleed H. Abdulla. 2024. "Advancing Ocular Imaging: A Hybrid Attention Mechanism-Based U-Net Model for Precise Segmentation of Sub-Retinal Layers in OCT Images" Bioengineering 11, no. 3: 240. https://doi.org/10.3390/bioengineering11030240

38. Inan, Yavuz Selim. "Adaptive multiscale retinal diagnosis: A hybrid trio-model approach for comprehensive fundus multi-disease detection leveraging transfer learning and siamese networks." arXiv preprint arXiv:2405.18449 (2024). https://doi.org/10.48550/arXiv.2405.18449
39. Abdullah, Ahmed Aizaldeen, Ahmed Aldhahab, and Hanaa Mohsin Al Abboodi. "Detection and Classification of Eye Diseases using Hybrid Deep Features with Decision Tree Algorithm." In 2024, the 3rd International Conference on Advances in Engineering Science and Technology (AEST), pp. 1-6. IEEE, 2024. 10.1109/AEST63017.2024.10960200
40. Al-Fahdawi, Shumoos, Alaa S. Al-Waisy, Diyar Qader Zeebaree, Rami Qahwaji, Hayder Natiq, Mazin Abed Mohammed, Jan Nedoma, Radek Martinek, and Muhammet Deveci. "Fundus-deepnet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images." Information Fusion 102 (2024): 102059. https://doi.org/10.1016/j.inffus.2023.102059
41. Mahmood, Salwa Shakir, Sihem Chaabouni, and Ahmed Fakhfakh. "A new technique for cataract eye disease diagnosis in deep learning." Periodicals of Engineering and Natural Sciences 11, no. 6 (2023): 14-26. https://doi.org/10.21533/pen.v11.i6.191
42. Fernandes, Veda. "Using data-efficient image transformers for diabetic retinopathy severity classification." Research Gate 6, no. 1 (2023): 7-9. https://doi.org/10.47611/harp.320
43. Shamsan, Ahlam, Ebrahim Mohammed Senan, and Hamzeh Salameh Ahmad Shatnawi. "Automatic classification of colour fundus images for prediction of eye disease types based on hybrid features." Diagnostics 13, no. 10 (2023): 1706. https://doi.org/10.3390/diagnostics13101706
44. He, Xiang, Yiming Wang, Fabio Poiesi, Weiye Song, Quanqing Xu, Zixuan Feng, and Yi Wan. "Exploiting multi-granularity visual features for retinal layer segmentation in human eyes." Frontiers in Bioengineering and Biotechnology 11 (2023): 1191803. https://doi.org/10.3389/fbioe.2023.1191803
45. Vázquez, Estefanía Rivas, María Noelia Barreira Rodríguez, Emilio López-Varela and Manuel G. Penedo. "Deep learning for segmentation of optic disc and retinal layers in peripapillary optical coherence tomography images." In Fifteenth International Conference on Machine Vision (ICMV 2022), vol. 12701, pp. 381-388. SPIE, 2023. https://doi.org/10.1117/12.2680545
46. Li, Zhenwei, Mengying Xu, Xiaoli Yang, and Yanqi Han. "Multi-label fundus image classification using attention mechanisms and feature fusion." Micromachines 13, no. 6 (2022): 947. https://doi.org/10.3390/mi13060947
47. https://www.kaggle.com/datasets/ftayush007/oia-odir-5k.
48. Abuya, Teresa Kwamboka, Richard Maina Rimiru, and George Onyango Okeyo. "An image denoising technique using wavelet-anisotropic Gaussian filter-based denoising convolutional neural network for CT images." Applied sciences 13, no. 21 (2023): 12069. https://doi.org/10.3390/app132112069
49. Kumar, Teerath, Rob Brennan, Alessandra Mileo, and Malika Bendechache. "Image data augmentation approaches: A comprehensive survey and future directions." Ieee Access 12 (2024): 187536-187571. https://doi.org/10.1109/ACCESS.2024.3470122
50. Kang, Kai, and Xiaogang Wang. "Fully convolutional neural networks for crowd segmentation." arXiv preprint arXiv:1411.4464 (2014). https://doi.org/10.48550/arXiv.1411.4464
51. Hussain, Hanan, P. S. Tamizharasan, and C. S. Rahul. "Design possibilities and challenges of DNN models: a review from the perspective of end devices." Artificial Intelligence Review (2022): 1-59 https://doi.org/10.1007/s10462-022-10138-z
52. Rokoni, Sakib. "LsEyeNet: deep feature fusion-based models for a large spectrum of eye disease recognition from ophthalmic images." PhD., BRAC University, 2025. http://hdl.handle.net/10361/26620
53. Pektaş, Muhammed. "Performance analysis of efficient deep learning models for multi-label classification of fundus image." Artificial Intelligence Theory and Applications 3, no. 2 (2023): 105-112. https://dergipark.org.tr/en/download/article-file/3202713?ref=https://git.chanpinqingbaoju.com