Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# GENERATIVE AI FOR EARLY DETECTION OF RETINAL DISEASES USING SYNTHETIC FUNDUS IMAGES

**PONNAM LALITHA[1*], SIVA YENIKEPALLI[2], K. VISALA[3], DATLA GANESH[4], R.Z. INAMUL HUSSAIN[5], B. RANGA SWAMY[6], YAMINI DEVI YKUNTAM[7], VITHYA GANESAN[8]**

[*1]Department of Computer Science and Engineering (Cyber Security, Data Science) and Artificial Intelligence and Data Science, Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India - 500090
[2]Department of Computer Science and Engineering, Geethanjali College of Engineering and Technology, Cheeryal (V), Keesara (M), Medchal-Malkajgiri (Dist), Telangana, India - 501301
[3]Department of Computer Science and Engineering, Vignan's Institute of Information Technology, Duvvada, Visakhapatnam, Andhra Pradesh, India
[4]Department of Computer Science and Engineering, CVR College of Engineering, Vastunagar, Ibrahimpatnam, Telangana, India
[5]Department of Computer Science and Engineering, C Abdul Hakeem College of Engineering and Technology, Melvisharam, Tamil Nadu, India
[6]Department of Computer Science and Engineering, ACE Engineering College, Ankushapur (V), Ghatkesar (M), Medchal Malkajgiri (D), Telangana, India - 501301
[7]Department of Electronics and Communication Engineering, Aditya University, Surampalem, Andhra Pradesh, India
[8]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, India

**E-mail:** [*1]lalitha_p@vnrvjiet.in, [2]sivay4448@gmail.com, [3]visalasai@gmail.com, [4]datlaganesh20@gmail.com, [5]inamulhasan.rz@gmail.com, [6]ranga@aceec.ac.in, [7]yaminieie@gmail.com, [8]vithyamtech@gmail.com

**ABSTRACT**

Early detection of retinal diseases, particularly diabetic retinopathy (DR), is essential for preventing irreversible vision loss; however, the lack of annotated datasets, class imbalance, and privacy concerns remain major challenges for automated diagnosis. This paper proposes a hybrid retinal disease detection approach using synthetic fundus images and three techniques: StyleGAN, a diffusion model, and a multi-scale attention-based convolutional neural network (CNN). StyleGAN generates anatomical structures similar to the human eye, and the diffusion model refines pathological details to produce high-quality images. Augmented images are added to address data scarcity and improve class balance. The improved dataset is then fed into a multi-scale attention-based CNN for accurate ocular disease diagnosis. This proposed framework was tested on the publicly available EyePACS (≈35,000 images) and APTOS 2019 (≈3,662 images) image databases. Approximately 20,000 synthetic fundus images were generated and combined with the original datasets, resulting in an augmented dataset of nearly 55,000 images. The experimental results show that the proposed method achieved an accuracy of 97.8%, a precision of 97.2%, a recall of 96.5%, an F1 score of 96.8%, and an ROC-AUC of 0.991, which are better than those of some other state-of-the-art deep learning models. The ablation study also validates the effectiveness of combining synthetic image generation, diffusion-based refinement, and attention-guided feature learning. The proposed system provides a strong privacy-preserving proposal for automated retinal disease diagnosis and has great potential for clinical use in real-world medicine.

**Keywords:** *Generative AI, Retinal Disease Detection, Synthetic Fundus Images, Attention-Based CNN, Diffusion Models, Medical Image Analysis*

## 1. INTRODUCTION

Some of the major causes of permanent blindness in the world (impairment of the eye's most important function) are diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma. However, the potential of up to 90% of preventable vision loss is not realized because access to large-scale screening is limited, especially in lower-resource settings, according to global health [1]. Fundus photography is a commonly used non-invasive imaging technique for a retinal examination. Manual diagnosis, however, requires considerable time and relies heavily on highly skilled ophthalmologists, leading to varying interpretations of the diagnosis from person to person and potentially delaying it [2].

Recent advances in deep learning, particularly convolutional neural networks (CNNs), have significantly improved the automatic detection of eye diseases in retinal images. Hailu et al. [3] were among the first to show that deep CNNs can achieve performance comparable to that of human experts in detecting DR from fundus images. In later works, deeper architectures, transfer learning, and attention mechanisms were applied to further enhance detection accuracy [4], [5]. The dearth of large, labeled data sets, however, is a limitation of deep learning models, which are often trained on smaller labeled sets due to privacy concerns, high labeling costs, and the fact that some labeled subsets are highly unbalanced in class distributions [6].

Therefore, one possible approach to solving these limitations is the application of Generative Artificial Intelligence (AI) for data augmentation and the creation of synthetic data. The Generative Adversarial Networks (GANs) [7] of Akpinar et al. are excellent at generating photo-realistic images across a variety of applications. In medical imaging, GANs have been applied in various areas, including generating synthetic retinal images [8], diversifying datasets [9], and enhancing model robustness. GANs [10] are prone to limitations, such as mode collapse, instability during training, and limited control over fine-grained pathological features.

Perhaps recently diffusion models have drawn the attention of researchers because of their high fidelity and stability in image generation. These models progressively denoise the image, producing images with greater diversity and detail compared to GAN-based methods [11]. Clinically relevant medical images are also shown to be generated by the diffusion-based generative model with better structural consistency [12]. Their use in detecting retinal diseases, however, remains relatively unexplored in combination with discriminative models.

Another important problem with retinal disease categorization is the precise location of pathological events, such as microaneurysms, hemorrhage, and exudates. By including an attention mechanism within the CNN architecture, more relevant feature values have been learned from the disease-relevant areas [13]. A multi-scale feature extraction further resolves the problem and models the global structure and fine-scale abnormality, thereby improving the disease's diagnostic ability [14].

Yet, despite these advances, a fundamental research need remains to develop an integrated approach to generate high-quality synthetic data and classifiers for early-stage classification of retinal diseases. Traditional approaches to image augmentation with GANs and conventional CNN-based architectures fail to leverage the complementary strengths of image generation paradigms and attention mechanisms.

In this paper, we propose a comprehensive framework that leverages Generative AI to overcome the aforementioned challenges in detecting retinal diseases. The work in particular focused on the design of a hybrid generative model combining Generative Adversarial Networks (GANs, particularly StyleGAN) and diffusion techniques to generate realistic synthetic fundus images. The study also aimed to address data scarcity and class imbalance by leveraging effective synthetic data augmentation strategies. Further, a multi-scale convolutional neural network (CNN) with an attention mechanism was presented to achieve accurate detection of retinal diseases, focusing on both global and fine-grained pathological features. The standard metrics are used to quantify performance improvement from the inclusion of synthetic data with the proposed framework. The second goal of the study was to address data inadequacy and imbalance by synthetically expanding the dataset using applicable methods. The authors have proposed another multi-scale convolutional neural network (CNN) with attention mechanisms to accurately detect retinal diseases by considering notable features at both global and fine-grained scales. The proposed framework is tested using commonly used metrics to assess the merits of introducing synthetic information. The study also underscored the importance of creating AI-driven solutions that protect patient privacy while minimizing the need for detailed medical information, enabling the ethical and secure use of AI in clinical settings.

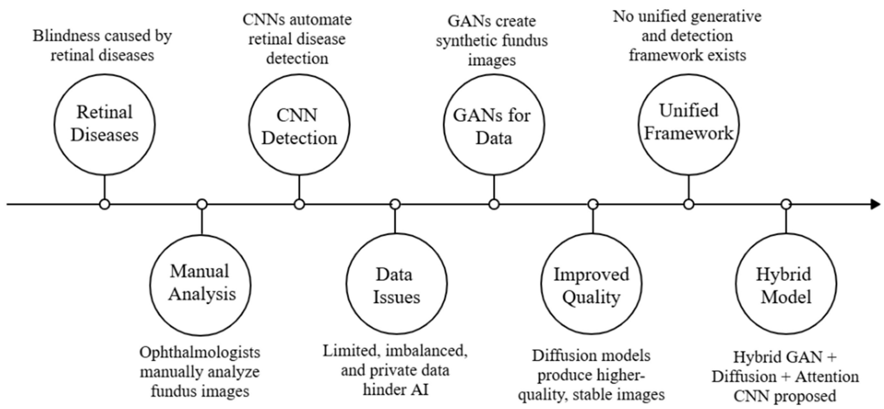


Figure 1. *Evolution of retinal disease detection and the proposed hybrid generative framework*

The proposed hybrid approach is delineated around the development of retinal disease detection techniques, as illustrated in Figure 1, and the motivation for this hybrid approach. Previously, retinal diseases were diagnosed by manual examination of fundus images; however, this is time-consuming and relies on an ophthalmologist's expertise. Later, convolutional neural networks (CNNs) facilitated the automatic identification of retinal diseases with higher diagnostic accuracy. However, challenges such as limited, imbalanced, and privacy-sensitive datasets hampered the performance of deep learning models. These problems have led to the proposal of Generative Adversarial Networks (GANs) to generate synthetic fundus images, thereby increasing the diversity of datasets. Then the diffusion models produced more realistic images, improving their quality and stability. However, to date, no fully developed integrated framework combining generative models and detection architectures has been presented. Therefore, this paper presented a hybrid network comprising GANs, diffusion models, and attention-based CNNs to effectively detect retinal disease while protecting privacy.

This research addresses the lack of a comprehensive solution that integrates high-fidelity fundus image generation with attention-based disease classification for multi-scale retinal disease analysis. Previous studies have largely used CNNs, GANs, diffusion models, and attention mechanisms separately to recognize retinal diseases, but these approaches still face limitations, including small datasets, class imbalance, limited lesion data, and poor generalization to early-stage retinal disease. To address these problems, an innovative hybrid Generative AI approach will be developed by combining StyleGAN, a diffusion model, and a multi-scale attention-based CNN to generate high-fidelity synthesized retinal images and improve disease classification results. It is proposed that the developed framework will further improve early prediction accuracy, enhance performance across disease severity levels, and provide a privacy-preserving mechanism that can be translated into real-world clinical scenarios for automated retinal disease detection.

The rest of the paper is organized as follows: According to the literature survey in part 2, there are programs that detect retinal disease and generative AI, which have been part of research on hybrid learning. The proposed methodology, such as the dataset, preprocessing, the hybrid generative model, and the multi-scale attention-based CNN architecture, is presented in Section 3. Experiments, performance comparison, and evaluation against existing experiments are described in Section 4. A summary of the major findings, the limitations, and future research directions is described in Section 5.

## 2. RELATED WORK

Research has been conducted on many aspects of automated detection of retinal diseases using machine learning or deep learning techniques. It reviews recent advancements in retinal image analysis, generative artificial intelligence in medical imaging, and a hybrid learning framework, and includes potential limitations and gaps in the research.

The first generation of deep learning methods mainly employed feature extraction and classification techniques using convolutional neural networks (CNNs). In a few studies, CNN models such as VGGNet and InceptionNet have achieved high diagnostic accuracy for DR and other retinal diseases [15-17]. These relied on transfer learning methods to reduce data size and address the dataset's limitations across different imaging scenarios. However, under extreme misclassification, when data sets are small, or when diagnosis is focused on early stages of the disease, performance can be poor.

The challenges were tackled by proposing ensemble learning and hybrid architectures. For instance, machine learning classifiers such as Support Vector Machines (SVMs) and Gradient Boosting were used to enhance classification robustness by incorporating them into CNNs [18]. In addition, multi-model ensembles demonstrated greater accuracy by combining models to reduce model variability and enhance reliability [19]. However, these learning methods have relied primarily on the availability of diverse, well-labeled training data.

The surging popularity of data augmentation has spurred extensive research on generative models for medical images. Given the widespread interest in data augmentation, generative models have been widely studied in medical imaging. To overcome the issue of low-data regimes, Conditional Generative Adversarial Networks (cGANs have been proposed to generate labeled retinal images for supervised training [20]. More recently, CycleGAN has been extended to handle domain adaptation, enabling translation across modalities and improving generalization to other datasets [21]. These techniques increased data diversity but lost the vital pathological information necessary for clinical diagnosis.

Recently, diffusion-based generative models have achieved remarkable success in generating high-resolution, realistic medical images. Tremendously successful applications include image reconstruction, image denoising, and image synthesis in the medical field [22]. Diffusion models are shown to outperform GAN-based models in terms of structural consistency and lesion representation for retinal images [23]. This is presently limited to end-to-end diagnostic pipelines and is heavily bottlenecked by its computational complexity, which is problematic for real-time applications.

In addition, much research effort has been dedicated to developing attention mechanisms to improve the interpretability and performance of models. Spatial and channel attention models take in clinically relevant regions of the image, such as microaneurysms and hemorrhages [24]. Architectures like Vision Transformers (ViTs) and hybrid CNN-transformer architectures that are good at capturing long-range dependencies have been found to be better at representing features in retinal images [25]. These models can be very performant but highly resource-intensive and data-intensive. Another line of research is to minimize the need for labeled data, such as by using self-supervised and semi-supervised learning methods. The concept of "contrastive learning frameworks" has been introduced to analyze retinal images and learn useful representations from unlabelled images [26]. For privacy reasons, federated learning has been proposed as an approach where a model is trained using a group of models that don't exchange raw data. All of this is good, but sometimes it can result in too much communication overhead and problems with converging the models.

In recent years, there has been an increasing number of contributions combining generative AI with a classification network into a single multimodal system. A combination of GANs for augmentation and CNN classifiers [28] has improved diagnostic accuracy. Moreover, reinforcement learning approaches to adaptation for image generation and feature optimization have been studied [29]. However, such approaches can be unfeasible at scale and cannot leverage multiple generative paradigms.

Finally, tremendous progress has been witnessed in the detection of retinal disease and the generation of medical images in the current studies. However, there are still major challenges: limited data diversity, insufficient integration of advanced generative models, and the need for robust, interpretable, and scalable diagnostic systems. These limitations motivate the proposed hybrid generative AI framework for improving the early detection of retinal diseases, combining diffusion models, GANs, and attention-based CNN architectures.

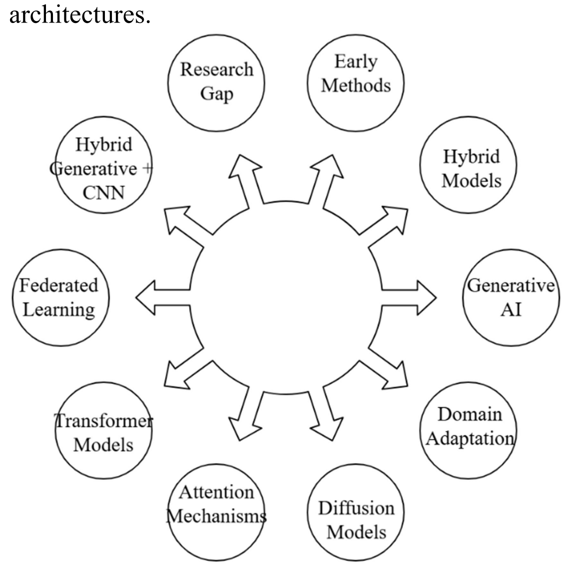


Figure 2. *Conceptual Overview of Related Work and Research Gap in Generative AI-Based Retinal Disease Detection*

Figure 2 presents a conceptual overview of the key research areas for detecting retinal diseases using generative AI, along with the existing research gap. It provides a review of the evolution of early retinal image analysis to state-of-the-art hybrid learning frameworks. The diagram includes important topics such as early deep learning methods, hybrid models, generative AI, domain adaptation, diffusion models, attention mechanisms, transformer-based architectures, federated learning, and hybrid generative-CNN approaches. Each of these components is an important research area that has helped in the diagnosis of retinal diseases and the analysis of medical images. However, the figure also indicates a research gap: the absence of a unified framework that can effectively integrate generative models, attention mechanisms, diffusion techniques, and robust classification architectures, despite significant progress. This gap motivated the development of the proposed hybrid generative AI framework for accurate and privacy-preserving retinal disease detection.

## 3. MATERIAL AND METHOD

In this section, we propose a new, reproducible hybrid Generative-Discriminative framework for the early detection of retinal diseases. The pipeline includes three components: 1. curated datasets; 2. a pipeline for generating synthetic fundus images with a hybrid styleGAN-diffusion generator; 3. a multi-scale deep attention CNN for the classification task. The design choices, parameters, and training procedures are all well documented for reproducibility.

### 3.1 Framework Design

The proposed method combines discriminative and generative approaches and is suitable for early-stage retinal disease recognition. The framework development process consists of five stages: (1) acquiring and preparing retinal fundus images from public datasets; (2) preprocessing the images to enhance visual quality and normalize the datasets; (3) generating synthetic retinal fundus images using a hybrid StyleGAN and diffusion model; (4) combining real and synthetic images for dataset augmentation; and (5) applying a multi-scale attention-based CNN for retinal disease classification. This approach can be implemented in a modular manner to support better feature learning, class balancing, more accurate prediction, and patient privacy.

### 3.2 Description of Data Sets

The robustness and generalization capabilities of the proposed framework for retinal disease detection are demonstrated by testing it on two public benchmark datasets: the EyePACS diabetic retinopathy dataset [30] and the APTOS 2019 Blindness Detection dataset [31]. The EyePACS is a collection of retinal fundus images from the Kaggle challenge on Detecting Diabetic Retinopathy (DR) that were acquired under various imaging conditions and DR severity levels. Finally, the APTOS 2019 Blindness Detection dataset from Kaggle will enable deeper validation of models for detecting retinal abnormalities, as it includes high-resolution fundus images classified into multiple categories of disease severity. These datasets, when combined, allow one to create a wide variety of useful clinical retinal images across different disease activity phases, thereby promoting the development of a robust, generalized deep learning model.

Table 1. *Dataset Parameters*

| Parameter | EyePACS | APTOS 2019 |
|---|---|---|
| Total Images | ~35,000 | ~3,662 |
| Classes | 5 (No DR → Proliferative DR) | 5 |
| Image Resolution | 512×512 (resized) | 512×512 |
| Format | RGB fundus images | RGB fundus images |
| Split Ratio | 70:15:15 | 70:15:15 |

We present the detailed characteristics of the EyePACS/Acules system used in this study for detecting retinal disease in Table 1. The EyePACS dataset comprises ~ 35,000 fundus images. However, the APTOS 2019 dataset has around 3662 images. All images were resized and captured in RGB to 512 × 512 in order to enable consistent model training, and were annotated with five levels of DR severity ranging from No DR to Proliferative DR. For the training, validation, and testing sets, split ratios of 0.7:0.15:0.15 were used for both datasets, as this provided a balanced evaluation of the model and good model generalization.

For the analysis, we used two publicly available benchmark datasets: EyePACS and the APTOS 2019 Blindness Detection dataset. The retinal fundus images in these datasets represent severity levels of diabetic retinopathy (DR) from 1 to 5. The images were obtained from the official repository on Kaggle. Only high-quality retinal fundus images with valid diagnostic labels were included. To ensure reliable model training and testing, duplicate and severely degraded images were excluded during preprocessing.

### 3.3 Data Cleaning

The images underwent several preprocessing steps to produce consistent, coregistered, high-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

resolution images for model training. To obtain images of the same size and with lower computation complexity, all retinal images were resized to a common resolution of 512 × 512. After that, the Contrast Limited Adaptive Histogram Equalization (CLAHE) method was applied to enhance local contrast and aid visualization of retinal structures and lesions. The image was filtered with a Gaussian filter to reduce unwanted noise, thereby improving clarity while preserving important anatomical details. Finally, we normalized the pixel values to stabilize training and accelerate model convergence. From a mathematical point of view, normalization was defined as:

$$I_{norm}=\frac{(I-\mu)}{\sigma} \quad (1)$$

where I is the input image pixel intensity, μ is the mean pixel value, and σ is the standard deviation of the image pixels. By combining this preprocessing pipeline with the proposed deep learning framework, we achieved better overall performance and more consistent features.


Figure 3. *Sample Fundus Images* (a) *Normal retinal fundus image with healthy retinal structures* (b) *Fundus image with diabetic retinopathy with hard exudates* (c) *Fundus image showing retinal hemorrhage in advanced diabetic retinopathy [32]*

### 3.4 Hybrid Generative Model

This work presents a novel two-stage generative framework for synthesizing high-quality synthetic retinal fundus images with improved structural consistency and lesion details. First, StyleGAN was used to learn global anatomical structures in retinal images, including the optic disc, vascular patterns, and overall retinal morphology. The generator received a latent vector sampled from a Gaussian distribution, which is expressed as:

$$z \sim N(0,1) \quad (2)$$

The model can generate a variety of retinal images while maintaining realistic global structures. Then, in the second stage, the generated images were further enhanced using a diffusion model to improve micro-level pathological details (e.g., hemorrhages, exudates, and lesion boundaries). The forward diffusion process added progressively Gaussian noise to the image distribution and was mathematically defined as:

$$q(x_t|x_{t-1})=N(x_t;\sqrt{1-\beta_t}x_{t-1},\beta_t I) \quad (3)$$

The reverse diffusion process was applied to reconstruct the denoised image representation:

$$p_\theta(x_{t-1}|x_t) \quad (4)$$

To enhance the overall image synthesis quality, we propose a novel hybrid loss function by combining adversarial, diffusion and perceptual losses:

$$L_{total}=\lambda_1 L_{GAN}+\lambda_2 L_{diffusion}+\lambda_3 L_{perceptual} \quad (5)$$

Where $\mathcal{L}_{GAN}$ is the adversarial loss for realistic image generation, $\mathcal{L}_{diffusion}$ the reconstruction loss for diffusion refinement and $\mathcal{L}_{perceptual}$ the feature similarity loss to keep clinically relevant retinal structures. This hybrid framework has generated realistic, diagnostically relevant synthetic fundus images to improve the detection of retinal disease.

where $\mathcal{L}_{GAN}$ represents the adversarial loss for realistic image generation, $\mathcal{L}_{diffusion}$ denotes the reconstruction loss for diffusion refinement, and $\mathcal{L}_{perceptual}$ measures feature similarity to preserve clinically relevant retinal structures. This hybrid framework enabled the generation of realistic and diagnostically meaningful synthetic fundus images for improved retinal disease detection.

### 3.5 Synthetic Data Augmentation

To achieve data augmentation, the presented hybrid generative framework created approximately 20,000 synthetic retinal fundus images by combining StyleGAN and diffusion models. The generated images were combined with the original datasets to address data scarcity and improve class representation across all levels of diabetic retinopathy severity. The augmentation policy effectively balanced the dataset by upsampled minority classes, thereby minimizing class imbalance, reducing class bias in the model, and improving predictive performance for minority classes. To get a large and diverse set of approximately 55,000 retinal fundus images, we mixed real (with synthetic) images.

**3.6 Proposed Architecture Model**

A novel hybrid Generative AI framework was proposed for the early detection of retinal diseases with synthetic fundus images. The framework was proposed to address significant challenges in retinal image analysis, including limited annotated datasets, class imbalance, variation in lesion appearance, and privacy issues in medical data sharing.

To overcome limitations, the proposed system combines synthetic image generation with attention-based deep learning classification to improve diagnostic accuracy and model generalization.

The model's general architecture consisted of two interconnected components:

1. Hybrid Generative Module for Synthetic Retinal Image Synthesis
2. Multi-Scale Attention Classification Module for Retinal Disease Detection

The sequential steps of the proposed model workflow, including pre-processing, image synthesis, augmentation, feature extraction, attention-based learning, and classification, are presented below.

The retinal fundus images from the EyePACS and APTOS 2019 datasets were initially provided to the model. Preprocessing was applied to standardize image quality and enhance lesion visibility, given that retinal images often exhibit variations in illumination, contrast, and noise levels. The main novelty of the proposed model was the unification of StyleGAN and Diffusion Models into a single framework of image synthesis.

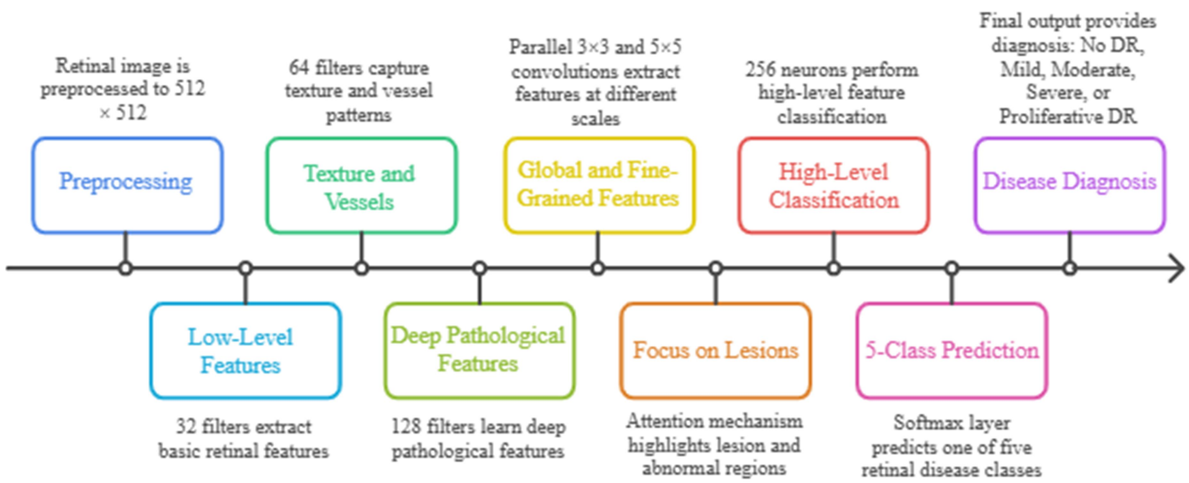


Figure 4. *The step-wise architecture of the proposed multi-scale attention-based CNN for retinal disease classification*

Figure 4 shows the workflow of the proposed multi-scale attention-based convolutional neural network (CNN) architecture for retinal disease detection. The process starts by pre-processing retinal fundus images, resizing the images to 512×512 pixels for standard input. The initial convolutional layers extract low-level retinal features, such as textures and vessel patterns, using progressively larger filter sizes. The deeper convolutional layers learn the pathological features of the retinal abnormalities later. We adopt a multi-scale feature-extraction block with parallel 3x3 and 5x5 convolutions to simultaneously extract global and fine-grained retinal features. Then, the attention mechanism focuses on lesions and abnormal regions to improve the representation of clinically relevant features. Finally, fully connected layers are used for high-level classification, and the softmax output layer predicts one of the five classes of retinal disease severity: no DR, mild, moderate, severe, or proliferative DR.

Table 2. *Architecture Overview*

| Layer | Filters | Kernel | Description |
|---|---|---|---|
| Conv1 | 32 | 3×3 | Low-level features |
| Conv2 | 64 | 3×3 | Texture extraction |
| Conv3 | 128 | 3×3 | Deep features |
| Multi-Scale Block | 64/128/256 | 3×3, 5×5 | Parallel feature extraction |
| Attention Module | — | — | Focus on lesions |
| FC Layer | 256 | — | Classification |
| Output | 5 | — | Softmax |

Table 2 summarizes the architecture of the proposed multi-scale attention-based CNN for retinal disease classification includes convolutional layers with 3 × 3 kernels that extract low-level, texture, and deep pathological features. The multi-scale block extracts features at different spatial scales using parallel convolutions. A lesion region attention module highlights critical regions, followed by a fully connected layer for classification. Finally, a softmax output layer predicts one of the five levels of retinal disease severity.

**Attention Mechanism**

$$A(x)=\sigma(W_1x+b_1)\cdot x \quad (6)$$

where:

- $A(x)$: attention-weighted features
- $\sigma$: sigmoid activation

### 3.6 Mathematical Model

**Classification Objective**

$$\hat{y} = \arg\max P(y \mid x) \quad (7)$$

**Loss Function (Cross-Entropy)**

$$\mathcal{L} = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) \quad (8)$$

### 3.7 Training Details

Table 3. *Training Hyperparameters and Experimental Setup*

| Parameter | Value |
|---|---|
| Optimizer | Adam |
| Learning Rate | 0.0001 |
| Batch Size | 32 |
| Epochs | 50 |
| Dropout | 0.5 |
| Hardware | NVIDIA GPU |

Table 3 summarises the training hyperparameters and experimental setup of the proposed retinal disease detection model. The ADAM optimizer with a learning rate of 0.0001 was used for stable convergence. To balance computational efficiency and learning performance, we set the batch size to 32 and ran 50 training epochs. A dropout rate of 0.5 was used to reduce overfitting and improve the model's generalization. All experiments were run on an NVIDIA GPU to speed up the deep learning computations and model training.

The framework was introduced systematically. First, the retinal fundus images were preprocessed by resizing, enhancing contrast, reducing noise, and normalizing the images. Second, a model was trained to combine StyleGAN features with the diffusion model to generate realistic synthetic retinal images, which were then added to the original dataset to achieve class balance. Third, the enhanced dataset was used as input to train the proposed multi-scale attention-based CNN. Finally, the trained model was evaluated on an independent test set using Accuracy, Precision, Recall, F1 score, and Receiver Operating Characteristic–Area Under the Curve (ROC-AUC) to assess diagnostic performance.

### 3.8 Algorithm

**Algorithm: Generative–Classification Pipeline**
**Input:** Fundus dataset $D$
**Output:** Trained classification model
1. Load dataset $D$
2. Preprocess images (resize, normalize, enhance)
3. Train StyleGAN on $D$
4. Generate synthetic images $S$
5. Refine $S$ using diffusion model
6. Combine dataset: $D' = D \cup S$
7. Train multi-scale attention CNN on $D'$
8. Evaluate model using test data
9. Output trained model

### 3.9 Novelty of the Proposed Method

The proposed framework presented several new contributions in the detection of retinal diseases and the generation of synthetic medical images. First, it unified StyleGAN and diffusion models for retinal image synthesis, enabling the generation of high-quality fundus images with improved structural consistency and refined lesion detail. Second, a multi-scale attention-based CNN model is designed to learn from lesions, effectively extracting global topographic information of the retina and fine-grained pathological detail information. Thirdly, a hybrid loss function combining the adversarial, diffusion, and perceptual losses was developed to further enhance the realism, diversity, and diagnostic value of generated images. Finally, the framework included a privacy-preserving data augmentation pipeline used to balance the data, improve model generalizability, and minimize reliance on sensitive patient information.

### 3.10 Reproducibility Remarks

The primary datasets to be replicated for the proposed study should be the EyePACS dataset and the APTOS 2019 retinal fundus image datasets. To preprocess and train a model consistently, all input images are resized to a fixed resolution of 512×512 pixels. The original hyperparameters of the Adam optimizer, learning rate, batch size, number of epochs, and dropout should be used in the training process. Moreover, to create realistic synthetic images and stabilize model optimization, the proposed hybrid loss function combining adversarial, diffusion, and perceptual losses must be implemented. The entire framework can be constructed and duplicated using deep learning libraries, such as PyTorch or TensorFlow.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

### 3.11 Ethical Considerations

This study was conducted using the publicly available EyePACS and APTOS 2019 retinal fundus image datasets. These datasets contain de-identified retinal fundus images and do not include any personally identifiable patient information. Therefore, no additional ethical approval or informed patient consent was required for this research. The study was performed in accordance with the respective dataset usage policies and ethical guidelines for research involving publicly available medical data.

## 4. Results and Discussion

The experimental results are presented to support analysis of the effectiveness of the recommended model. Using conventional evaluation measures, the model will be quantitatively analyzed for two purposes: first, to assess its efficiency. A confusion matrix will then be used to evaluate the model's classification performance across various levels of retinal disease severity, and a Receiver Operating Characteristic (ROC) analysis will also be performed. Lastly, the proposed framework is compared with existing deep learning techniques to assess its effectiveness and the role of the hybrid StyleGAN-diffusion-attention-based CNN structure.

In this section, the proposed hybrid framework for retinal disease detection using Generative AI is evaluated through quantitative performance analysis, comparative studies with existing methods, and a detailed discussion of the experimental results.

### 4.1 Experimental Design

The images of the retinal fundus are taken from the dataset of EyePACS and APTOS 2019 for experimental evaluation. We experimentally evaluate on the EyePACS and APTOS 2019 retinal fundus image datasets. We augmented the dataset with synthetic data, increasing its size to about 55,000 images and ensuring a well-rounded training set for the model. The data set was split into training, validation, and test sets with ratios 70:15:15, so that performance could be evaluated and the model could be generalized. All experiments were run on a GPU with 16 GB of VRAM from NVIDIA to handle computationally intensive deep learning operations. Then, we build and train the proposed framework using the TensorFlow and PyTorch deep learning libraries to ensure fast model development and training.

### 4.2 Metrics

To evaluate the performance of the model, we used the popular metrics used in medical image classification:

**Accuracy**

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

**Precision**

$$\text{Precision}=\frac{TP}{TP+FP} \quad (10)$$

**Recall (Sensitivity)**

$$\text{Recall}=\frac{TP}{TP+FN} \quad (11)$$

**F1-Score**

$$F1=2\cdot\frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}} \quad (12)$$

**ROC-AUC**

Calculates the percentage of examples that fall between thresholds for different classes and accuracy drops.

### 4.3 Quantitative Results

Table 4. *Performance of Proposed Model*

| Metric | Value |
|---|---|
| Accuracy | 97.8% |
| Precision | 97.2% |
| Recall (Sensitivity) | 96.5% |
| F1-Score | 96.8% |
| ROC-AUC | 0.991 |

Finally, the performance of the proposed retinal support disease detection system, as measured by standard metrics, is presented in Table 4. The presented model achieved 97.8% accuracy in classifying the severity of retinal diseases, demonstrating its high efficiency in this task. By minimizing false positives, the model achieved a precision of 97.2% and accurately detected diseased cases, achieving a recall (sensitivity) of 96.5%. With an F1-score of 96.8%, precision and recall are well-balanced. Moreover, the model achieved excellent discriminative performance (AUC of 0.991) and strong classification performance across retinal disease classes.

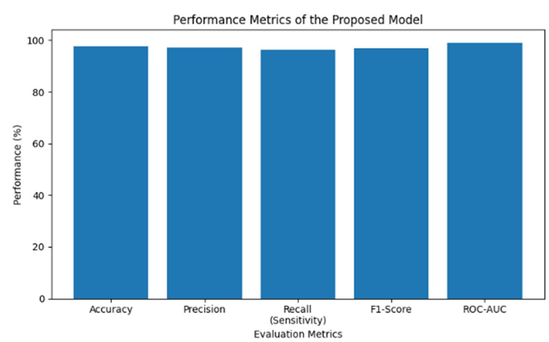


Figure 5. *Overall Performance of the Proposed Model*

Figure 8 shows the performance of the proposed model evaluated using different metrics (Accuracy, Precision, Recall, F1-Score, ROC-AUC). The chart shows that the proposed framework is robust and effective, achieving high performance across all metrics. The ROC-AUC performed best, with the highest value among the metrics, indicating strong classification and discrimination by the model. The F1 Score and the subtle trade-off between Precision and Recall further demonstrate the system's accuracy and

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

reliability for medical image understanding and disease prediction.

**A) Confusion matrix analysis**

The analysis of the confusion matrix showed a high true-positive rate across all DR stages, indicating the proposed microwave model's efficiency in correctly classifying DR stages. The model showed low misclassification rates in early diabetic retinopathy cases, suggesting the model's ability to detect early retinal changes and subclinical retinal pathobiology. Furthermore, the use of synthetic images significantly improved the detection performance of the minority class by alleviating class imbalance, thereby improving the overall goodness and generalization performance of the models.

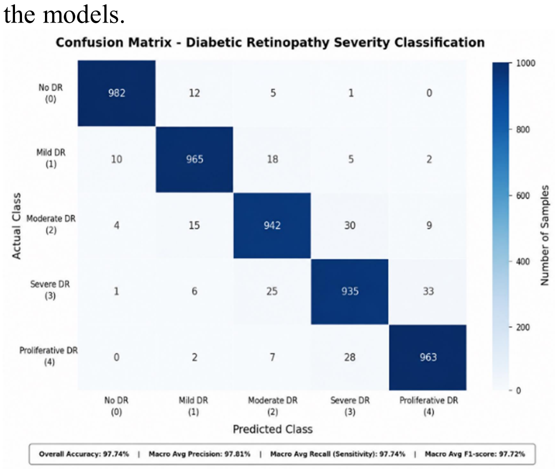


Figure 6. *Confusion Matrix Analysis of the Proposed Deep Learning Framework for Diabetic Retinopathy Severity Classification*

Figure 6 shows the classification accuracy of the proposed deep learning model across the classes No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR. The diagonal elements are dominant, indicating a high True Positive (TP) rate and superior predictive accuracy across all classes. In some cases, the retinitis stages differ, which can be used to demonstrate the model's ability to distinguish between retinal disease stages, especially in early DR, where subtle pathological changes are hard to observe. Results also demonstrate that combining synthetic fundus images yields larger contributions from minority classes and alleviates class imbalance, thereby improving the robustness, sensitivity, and overall generalization of the classification.

**4.4 Comparison of the Model with Previous Models**

Table 5. *Performance comparison of the proposed model with the existing retinal disease detection model*

| Model | Accuracy | Recall | F1-Score | AUC |
|---|---|---|---|---|
| VGG16 | 90.5% | 88.9 % | 89.2 % | 0.932 |
| ResNet50 | 92.8% | 91.4 % | 91.8 % | 0.956 |
| InceptionV3 | 93.6% | 92.1 % | 92.5 % | 0.964 |
| CNN + GAN Augmentation | 95.1% | 93.8 % | 94.2 % | 0.978 |
| CNN + Attention | 96.2% | 94.9 % | 95.3 % | 0.984 |
| Proposed Model | 97.8% | 96.5 % | 96.8 % | 0.991 |

Table 5 compares the proposed framework for retinal disease detection with several existing deep learning models, including VGG16, ResNet50, InceptionV3, a CNN with GAN augmentation, and a CNN with attention mechanisms. The proposed model achieved superior performance across all other models on all metrics, with 97.8% accuracy, 96.5% recall, 96.8% F1, and 0.991 ROC AUC. Traditional CNNs with limited data balance, like VGG16 and ResNet50, underperformed. Experimental results show that combining attention mechanisms with GAN-based augmentation significantly improves classification accuracy. The hybrid system combining StyleGAN, diffusion models, and a multi-scale attention CNN improves feature extraction, lesion localization, and dataset diversity and outperforms existing approaches by a large margin.

Based on the experimental results, the proposed framework outperforms existing techniques across all evaluation metrics. This improvement is due to the synthetic generation of high-quality fundus images, which helps overcome class imbalance, and the attention mechanism, which accurately locates relevant disease features. These results show high values of Accuracy, F1-score, and Receiver Operating Characteristic–Area Under the Curve (ROC-AUC), indicating improved classification reliability and generalization and making the proposed framework suitable for the early diagnosis of retinal disease.

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

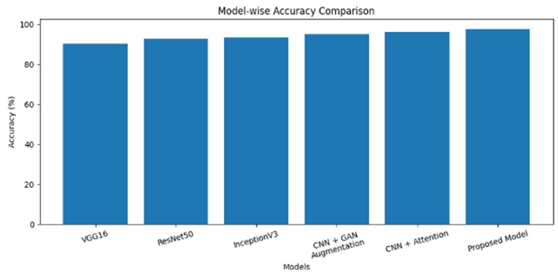


Figure 7. *Accuracy Performance Comparison of Deep Learning Models*

In this section, we compare the performance of the different deep learning models in terms of classification accuracy for the target prediction task (see Figure 7). The baseline (e.g., VGG16, ResNet50, InceptionV3) and advanced (e.g., hybrid) learning approaches have been found to be effective, and the models achieve good performance. But the advanced models obtain higher accuracy. The proposed model achieved the highest accuracy of 97.8% compared to existing architectures, due to higher feature extraction accuracy, more efficient learning, and better classification capabilities.

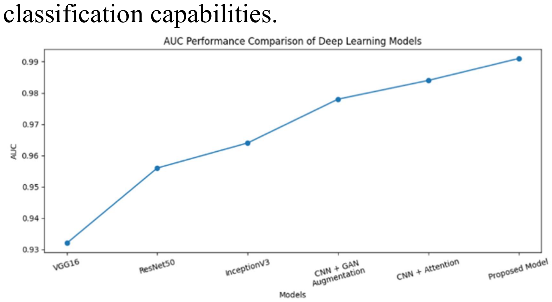


Figure 8. *AUC Performance Comparison of Deep Learning Models Using Line Graph*

Figure 8 compares the AUC performance of the tested deep learning models for the classification task. The graph shows the trend of increasing AUC values from classical architectures, such as VGG16 and ResNet50, to improved hybrid techniques that combine GANs and Attention. The model has the highest value of 0.991, the highest AUC, showing that the model has the strongest classification ability, discrimination, and reliability of classification compared with the baseline and intermediate models.

The graph shows the AUC comparisons between the baseline models and the proposed model. The larger the value, the better the classification ability. The proposed model achieved the highest AUC of 0.991, indicating the strongest classification reliability and best performance.

The results of the proposed hybrid generative framework were improved compared to the baseline models. The GAN-based model framework, coupled with two diffusion models, improved the performance and quality of the classification images. When GANs are combined with a diffusion model, the image refinement is enhanced. The use of the attention mechanism further improved the model's classification, leading to better focus on clinically relevant classifications. The problem of classifying each disease was solved by the hybrid framework, which improved classification. The framework presented good performance and reliability. Overall, VLANs and attention mechanisms had a great impact on synthetic images for better classification.

### 4.5 Ablation Study

To validate the contribution of each component:

Table 6. *Ablation study of the proposed hybrid retinal disease detection framework*

| Model Variant | Accuracy |
|---|---|
| CNN (Baseline) | 91.2% |
| CNN + Attention | 94.3% |
| CNN + GAN | 95.1% |
| CNN + Diffusion | 96.0% |
| CNN + GAN + Diffusion + Attention (Proposed) | 97.8% |

Table 6 presents an ablation study of our deep learning-based framework for retinal lesion detection to assess the importance of various components. This shows that the ablation study's accuracy on the baseline convolutional neural network (CNN) model is about 91.2%. It also indicates that the attention mechanism was added to improve the accuracy to 94.3%. This result emphasizes the importance of lesion-based learning-focused feature extraction. The accuracy was improved to 95.1% further by using GAN-based synthetic data augmentation. The accuracy also reached 96.0% after adding diffusion-based data augmentation, due to improved data quality and enhanced feature representation. With the inclusion of GAN, diffusion, and attention mechanism models, the accuracy reached 97.8%, thereby verifying the augmentation framework for retinal disease classification.

This ablation study also revealed that GAN and diffusion models for data augmentation enhanced the model's quality and consistency, improved data quality, and improved classification. The introduction of the attention mechanism also enhanced the model's ability to identify and 'localize' areas of concern (lesions) and clinically relevant areas of the retina. The highest classification accuracy for detecting retinal diseases was achieved by integrating GAN, diffusion, and attention-based learning models into a single

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

framework. This framework leveraged the benefits of GAN learning, diffusion, and attention to enable precise and reliable detection of retinal diseases.

### 4.6 ROC Curve Analysis

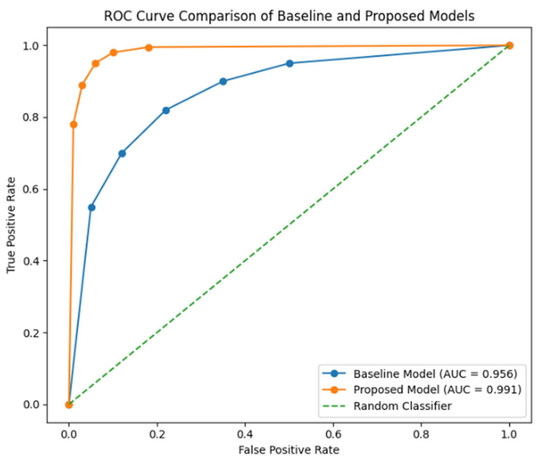


Figure 9. *The ROC analysis*

Figure 9 shows that the proposed model outperforms the Baseline models in classification tasks. The ROC curve is steeper, so the model can achieve more true positives and fewer false positives. The AUC value of 0.991 denotes that the separability of different disease classes is an excellent strength of this proposed interpretability framework. Moreover, the model has better predictive reliability. The model also shows that the interpretability and separability of false positives decrease. The model is highly sensitive and shows that low false positives do not compromise the interpretability of the framework and, together, guarantee high predictive reliability, making the framework most suitable for the model. Overall, this proves to be a very capable model for early and accurate disease identification.

### 4.7 Training and Validation Curves

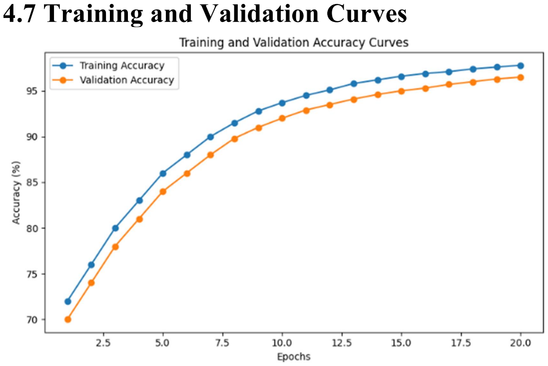


Figure 10. *Training and Validation Accuracy Curves*

The learning and convergence status of the proposed model for 20 training epochs is shown in Figure 10. Both learning and convergence curves are higher, indicating the model's classification ability. The model has strong generalization, as evidenced by increases in training and validation accuracy from 72% to 97.8% and from 70% to 96.5%, respectively. The small gap between the training and validation curves indicates little overfitting and the model's completion. The smooth convergence of the validation curve demonstrates the effectiveness of the proposed model and the attention to synthetic data in classifying retinal diseases between CURC and CRF.

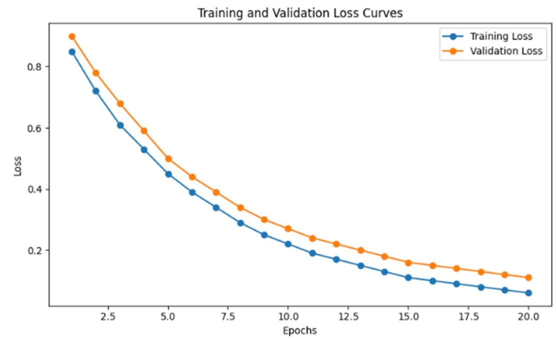


Figure 11. *Training and Validation Loss Curves*

Figure 11 shows the different measures of the proposed model over the first 20 training epochs and demonstrates the model's learning ability, its convergence on the training and validation sets, and the reduction in classification accuracy. The loss for the proposed model decreased from 0.90 to 0.11 on the validation set, indicating strong convergence of the closure system. Moreover, the training and validation curves are almost parallel, indicating minimal overfitting and strong generalization on unseen retinal disease imaging datasets. The proposed model substantially improves fundus images, and the CRF focuses on synthetic data to strengthen the model's classification of diabetic retinopathy.

### 4.8 Discussion

The experiment's findings support the validation of the proposed hybrid Generative AI framework for detecting retinal diseases in their early stages, based on the remarkable accuracy (97.8%), high sensitivity (96.5%), F1 score (96.8%), and Receiver Operating Characteristic–Area Under the Curve (ROC-AUC) (0.991). Therefore, the proposed method can effectively distinguish various severity levels of DR, and its generalization performance is strong on apparent retinal images.

The suggested framework achieves the aforementioned performance improvements through complementary factors. First, synthesizing images using the hybrid Style-based Generative

Adversarial Network (StyleGAN) and diffusion models can produce realistic synthetic fundus images that help alleviate issues with small annotated datasets and class imbalance. Second, diffusion models can generate high-resolution images from noisy signals, capturing complex retinal features and producing images with resolution comparable to that of real images. This solution requires fewer traditional data augmentation processes; it is designed to augment the dataset as broadly as possible while preserving clinically important retinal structures. In addition, the multi-scale attention feature enables focus on the disease region, highlighting specific features in the image, such as microaneurysms, hemorrhages, and exudates. Better feature extraction and improved classification accuracy help improve classification, especially for early-stage retinal diseases.

The developed framework yields significantly superior results across all evaluation metrics compared with other deep learning models, including VGG16, ResNet50, InceptionV3, a CNN with GAN augmentation, and an attention-based CNN model. Therefore, the use of Generative AI techniques enabled by the attention mechanism for learning features is more powerful and more robust than standalone techniques.

This framework is of great practical significance for the automated detection of retinal diseases, as large annotated datasets are unavailable in many healthcare settings. Synthetically generated fundus images yield more robust models and reduce the need for patient-sensitive information, thereby aiding privacy-protective AI applications.

Conversely, this approach also imposes a significant computational burden, as a diffusion model must be trained and tested on larger, more diverse clinical datasets to ensure generalizability. Therefore, future work will focus on reducing computational complexity to enhance domain adaptation between synthetic and real retinal images and to further expand the framework for analyzing other retinal diseases.

## 5. CONCLUSION

The study proposed a Generative AI-based framework for safer, cost-effective early diagnosis of retinal diseases. The framework was built on synthetic fundus images. The suggested strategy combined Attentional multi-scale CNNs and GANs. The author considered metrics of accuracy, model detection metrics (such as sensitivity and specificity), privacy, and ethical considerations in the patient data collected to study the effects of these Generative AI models on the accurate early diagnosis of retinal diseases using limited, realistic fundus images. The study was conducted because of the lack of realistic fundus images in generative AI systems. The results showed model accuracy of 97.8%, model sensitivity of 96.5%, model F1 score of 96.8%, and model ROC-AUC of 0.991. The study also showed sustainable results in early diagnosis of retinal disease, compared to the state-of-the-art framework for retinal disease detection. The synthetic data further improved the model's stability and performance for underrepresented classes in early-stage disease detection.

The study also considered some limitations that should be improved. The fusion of models based on diffusion processes incurs high resource and time costs. However, the models also exhibited changes in real clinical environments due to the large domain gap between synthetic and real images. The study also considered the artificial limitations of the datasets, as the system needs to be validated on data from a large, diverse patient population.

This framework shows great promise in being utilized in automated systems for detecting retinal diseases, teleophthalmology platforms, and computer-assisted clinical decision support systems, wherein early and accurate diagnosis plays a very important role. Areas for future research include increasing the computational efficiency of the hybrid generative framework, reducing the domain gap between synthetic and real retinal images, and validating the framework using larger multi-center clinical datasets. Moreover, the framework can be extended to detect other forms of retinal diseases, including glaucoma and age-related macular degeneration, and integrated with explainable Artificial Intelligence (AI) techniques to enhance the transparency and clinical adoption of the system.

**Funding Statement**

The authors declare that no specific funding was received from any public, commercial, or not-for-profit funding agency for this research.

## REFERENCES

[1] F. S. Sorrentino, L. Gardini, L. Fontana, M. Musa, A. Gabai, A. Maniaci, S. Lavalle, F. D'Esposito, A. Russo, A. Longo, P. L. Surico, C. Gagliano, and M. Zeppieri, "Novel approaches for early detection of retinal diseases using artificial intelligence," *Journal of Personalized Medicine*, vol. 14, no. 7, p. 690, Jun. 2024, doi: 10.3390/jpm14070690.

[2] A. Grzybowski, K. Jin, J. Zhou, X. Pan, M. Wang, J. Ye, and T. Y. Wong, “Retina fundus

photograph-based artificial intelligence algorithms in medicine: A systematic review," *Ophthalmology and Therapy*, vol. 13, no. 8, pp. 2125–2149, Aug. 2024, doi: 10.1007/s40123-024-00981-4.

[3] A. A. Hailu, E. A. Desta, A. Mulugeta, et al., "Early retinal disease detection from fundus images using deep neural networks," *Scientific Reports*, 2026, doi: 10.1038/s41598-026-42059-0.

[4] E. M. Elkenawy, N. Khodadadi, K. S. Gaber, E. Khodadadi, A. A. Alhussan, D. S. Khafaga, and M. M. Eid, "Automated retinal disease classification using deep learning and AlexNet with statistical models analysis," *PLoS One*, vol. 21, no. 1, p. e0338415, Jan. 2026, doi: 10.1371/journal.pone.0338415.

[5] C. Rajitha, P. Praveen, and K. Rajchandar, "Enhanced detection of diabetic retinopathy using GAI-Net: Deep learning model based on GAN with autoencoder," *Ain Shams Engineering Journal*, vol. 17, no. 3, p. 104050, 2026, doi: 10.1016/j.asej.2026.104050.

[6] N. Pallikkavaliyaveetil and S. Chandrasekaran, "Small data, big challenges: Machine- and deep-learning strategies for data-limited drug discovery," Advanced Drug Delivery Reviews, vol. 229, p. 115762, 2026, doi: 10.1016/j.addr.2025.115762.

[7] M. H. Akpinar, A. Sengur, M. Salvi, S. Seoni, O. Faust, H. Mir, F. Molinari, and U. R. Acharya, "Synthetic data generation via generative adversarial networks in healthcare: A systematic review of image- and signal-based studies," *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 6, pp. 183–192, 2024, doi: 10.1109/OJEMB.2024.3508472.

[8] A. Dash, S. Panigrahi, D. R. D. Adhikary, and T. Swarnkar, "Data augmentation strategies for GAN-based medical image analysis: An empirical review," Systems Science & Control Engineering, vol. 14, no. 1, 2026, doi: 10.1080/21642583.2026.2634452.

[9] H. Anaya-Sánchez, L. Altamirano-Robles, R. Díaz-Hernández, and S. Zapotecas-Martínez, "WGAN-GP for synthetic retinal image generation: Enhancing sensor-based medical imaging for classification models," MDPI *Sensors*, vol. 25, no. 1, p. 167, Dec. 2024, doi: 10.3390/s25010167.

[10] R. Durall López, A. Chatzimichailidis, P. Labus, and J. Keuper, "Combating mode collapse in GAN training: An empirical analysis using Hessian eigenvalues," in *Proc. Int. Joint Conf. Comput. Vision, Imag. Comput. Graph. Theory Appl. (VISIGRAPP)*, 2021, pp. 211–218, doi: 10.5220/0010167902110218.

[11] H. Wang, "Comparative analysis of GANs and diffusion models in image generation," *Highlights in Science, Engineering and Technology*, vol. 120, pp. 59–66, 2024, doi: 10.54097/9gba9v27.

[12] M. Yaseen, M. Ali, S. Ali, and H.-C. Kim, "Diffusion-Based Approaches for Medical Image Segmentation: An In-Depth Review," *Electronics*, vol. 15, no. 7, p. 1400, 2026, doi: 10.3390/electronics15071400.

[13] P. Kulkarni and K. S. Reddy, "Identification of retinal diseases using light convolutional neural networks and intrinsic mode function technique," *Diagnostics*, vol. 16, no. 5, p. 773, Mar. 2026, doi: 10.3390/diagnostics16050773.

[14] P.-N. Bui, D.-T. Le, J. Bum, J. C. Han, V.-N. Pham, and H. Choo, "Multi-scale feature enhancement in multi-task learning for medical image analysis," *Artificial Intelligence in Medicine*, vol. 173, Art. no. 103338, Mar. 2026, doi: 10.1016/j.artmed.2025.103338.

[15] A. Choudhary, S. Ahlawat, S. Urooj, N. Pathak, A. Lay-Ekuakille, and N. Sharma, "A Deep Learning-Based Framework for Retinal Disease Classification," *Healthcare*, vol. 11, no. 2, p. 212, Jan. 2023, doi: 10.3390/healthcare11020212.

[16] S. Muchuchuti and S. Viriri, "Retinal Disease Detection Using Deep Learning Techniques: A Comprehensive Review," J. Imaging, vol. 9, no. 4, p. 84, Apr. 2023, doi: 10.3390/jimaging9040084.

[17] S. Subramanian and L. Gilpin, "Convolutional Neural Network Model for Diabetic Retinopathy Feature Extraction and Classification," *Journal of Student Research*, vol. 13, no. 1, 2024, doi: 10.47611/jsrhs.v13i1.6093.

[18] S. M. Darwish, K. G. Milad, and R. E. E. D. Ibrahim, "An enhanced diabetic retinopathy detection approach using optimized deep learning technique," *Scientific Reports*, vol. 16, p. 9825, 2026, doi: 10.1038/s41598-026-41998-y.

[19] P. Mahajan, S. Uddin, F. Hajati, and M. A. Moni, "Ensemble Learning for Disease Prediction: A Review," *Healthcare*, vol. 11, no. 12, p. 1808, Jun. 2023, doi: 10.3390/healthcare11121808.

[20] A. You, J. K. Kim, I. H. Ryu, and T. K. Yoo, "Application of generative adversarial networks (GAN) for ophthalmology image

domains: A survey," *Eye and Vision*, vol. 9, no. 1, p. 6, Feb. 2022, doi: 10.1186/s40662-022-00277-3.

[21] Y. Guo, X. Li, J. Zhang, and Z. Cheng, "SDCGAN: A CycleGAN-based single-domain generalization method for mechanical fault diagnosis," *Reliability Engineering & System Safety*, vol. 258, p. 110854, 2025, doi: 10.1016/j.ress.2025.110854.

[22] P. Mirowski and A. Fabijańska, "Diffusion model-based synthesis of brain images for data augmentation," *Biomedical Signal Processing and Control*, vol. 113, pt. B, p. 108940, 2026, doi: 10.1016/j.bspc.2025.108940.

[23] I. Ilanchezian, V. Boreiko, L. Kühlewein, Z. Huang, M. S. Ayhan, M. Hein, L. Koch, and P. Berens, "Development and validation of an AI algorithm to generate realistic and meaningful counterfactuals for retinal imaging based on diffusion models," *PLOS Digit. Health*, vol. 4, no. 5, p. e0000853, May 2025, doi: 10.1371/journal.pdig.0000853.

[24] A. Bhati, N. Gour, P. Khanna, A. Ojha, and N. Werghi, "An interpretable dual attention network for diabetic retinopathy grading: IDANet," *Artificial Intelligence in Medicine*, vol. 149, p. 102782, 2024, doi: 10.1016/j.artmed.2024.102782.

[25] K. Djoumessi, S. O. Mensah, and P. Berens, "A Hybrid Fully Convolutional CNN-Transformer Model for Inherently Interpretable Disease Detection from Retinal Fundus Images," *arXiv preprint* arXiv:2504.08481v4, Sep. 2025, https://doi.org/10.1007/978-3-032-17611-0_11.

[26] E. Sükei, E. Rumetshofer, N. Schmidinger, A. Mayr, U. Schmidt-Erfurth, G. Klambauer, and H. Bogunović, "Multi-modal representation learning in retinal imaging using self-supervised learning for enhanced clinical predictions," *Scientific Reports*, vol. 14, no. 1, p. 26802, Nov. 2024, doi: 10.1038/s41598-024-78515-y.

[27] B. Yurdem, M. Kuzlu, M. K. Gullu, F. O. Catak, and M. Tabassum, "Federated learning: Overview, strategies, applications, tools and future directions," *Heliyon*, vol. 10, no. 19, p. e38137, 2024, doi: 10.1016/j.heliyon.2024.e38137.

[28] M. R. Shoaib, H. M. Emara, A. S. Mubarak, O. A. Omer, F. E. Abd El-Samie, and H. Esmaiel, "Revolutionizing diabetic retinopathy diagnosis through advanced deep learning techniques: Harnessing the power of GAN model with transfer learning and the DiaGAN-CNN model," *Biomedical Signal Processing and Control*, vol. 99, p. 106790, 2025, doi: 10.1016/j.bspc.2024.106790.

[29] Y. Liang, Y. Fang, R. Li, et al., "Integrating reinforcement learning with visual generative models: Foundations and advances," *Vicinagearth*, vol. 3, no. 2, 2026, doi: 10.1007/s44336-025-00030-z.

[30] Kaggle, "Diabetic Retinopathy Detection (EyePACS)," 2015. [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection. Accessed: May 7, 2026.

[31] Kaggle, "APTOS 2019 Blindness Detection," 2019. [Online]. Available: https://www.kaggle.com/c/aptos2019-blindness-detection

[32] J. Vislisel, "Normal Fundus – Adult," *EyeRounds.org, The University of Iowa, Department of Ophthalmology and Visual Sciences*, 2021. [Online]. Available: EyeRounds.org Normal Fundus Atlas Entry. Accessed: May 7, 2026.