Hybrid CNN-LSTM Architecture for OCT Retinal Disease Classification

Ehsan Khodadadi[1,*], P. K. Dutta[2], Amel Ali Alhussan[3], Marawa Metwally[4,5]

[1]Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701, USA
[2]School of Engineering and Technology, Amity University Kolkata, Kolkata, India
[3]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[4]Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA
[5]Jadara Research Center, Jadara University, Irbid 21110, Jordan

Emails: Ehsank@uark.edu; pkdutta@kol.amity.edu; aaalhussan@pnu.edu.sa; marwa@jcsis.org

Abstract

The ability to accurately classify retinal fundus images has been made possible by rapid improvements in deep learning (DL) and artificial intelligence (AI). This motivation led to developing a new AI-driven hybrid Convolutional Neural Network with Long Short-Term Memory (CNN-LSTM) architecture for precisely categorizing retinal diseases. The model first receives high-resolution retinal fundus images to extract various spatial properties, which are then processed by two parallel CNN branches after a standard convolutional layer. These branches use residual learning with convolutional and identity blocks to extract features. Following the reshaping and concatenation of the features from both branches, an LSTM layer captures inter-feature relationships. Eight retinal disorders are then predicted to belong to the same disease class via a fully linked classifier. Extensive simulations were run on a benchmark retinal OCT dataset, and performance was assessed using various criteria. The experimental results showed that the suggested hybrid model was adequate, with a high overall accuracy of 93% with F1-score values of 0.93, 0.94, and 0.93 for precision, recall, and accuracy, respectively. The model demonstrated considerable predictive abilities for all classes while perfectly classifying AMD, CNV, CSR, DME, DR, MH, and routine diseases to reveal its clinical value as an automated retinal processor.

Keywords: Retinal disease classification; Deep learning; CNN-LSTM architecture; Fundus imaging; Residual learning; Optical Coherence Tomography (OCT); Medical image analysis

1 Introduction

The two most prevalent instances of a life-endangering eye disease in terms of baseline eye health among the global population are diabetic retinopathy (DR) and age-related macular degeneration (AMD). Optical coherence tomography (OCT) usage in retinal diagnostics has changed non-invasive cross-sectional retinal observation to the higher detail examination with the help of advanced imaging methods [1]. However, the contemporary diagnostic methods based on OCT technology are characterized by critical limitations in terms

DOI: https://doi.org/10.54216/JAIM.100201
Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

of the consistency with the high level of accuracy in terms of different clinical settings and different patient groups [2]. The variations in the quality of images, the difference of the disease progression, and the presence of both the pathological manifestations creates invalid diagnostic results [3].

Diabetic retinopathy is a microvascular complication of diabetes, affecting small blood vessels in the retina through pericyte loss and endothelial cell deterioration, ultimately leading to increased capillary permeability [4]. The condition progresses gradually and remains asymptomatic in early phases, requiring routine ophthalmologic examinations to ensure timely recognition of the condition. Although manual reading of retinal photographs is comparatively more accurate than conventional dilated eye examinations, it represents one of the major screening modalities in clinical practice [5]. One of the most important factors in maintaining vision among patients with long-term diabetes is early detection followed by timely intervention [6].

In AMD, there is the development of drusen deposits, lipofuscin-containing material, and phospholipid vesicles between the retinal pigment epithelium basement membrane and Bruch's membrane [7]. Such structural modifications are also a source of significant global visual impairment among the older generations [8]. Proper identification of retinopathy and maculopathy necessitates professional skills and considerable clinical experience, which is a problem in resource-limited environments [9].

There are significant challenges encountered by healthcare systems related to large-scale deployment of deep learning-based medical image classification due to insufficient annotated medical cases to produce standardized datasets [10]. Transfer learning provides a convenient solution in the following manner: diagnostic models can apply the knowledge of pre-trained networks and, in this way, extend their diagnostic abilities [11]. In this study, transfer learning approaches are employed to design an automated computer-aided diagnosis tool that is able to identify AMD, choroidal neovascularization (CNV), DR, diabetic macular edema (DME), central serous retinopathy (CSR), drusen, macular hole (MH), and normal retinal conditions [12].

The artificial intelligence (AI) has become a common solution among ophthalmology practitioners that want to leverage the efficacy and precision of diagnostic tools and incorporate them into clinical practices and procedures [13], [14]. Deep learning, as a specialized subset of AI, has shown significant potential in medical image analysis, in that it is capable of automatically extracting relevant features without the need for human intervention [15]. Convolutional neural networks (CNNs) have demonstrated outstanding performance in medical imaging tasks, including tumor detection, organ segmentation, and disease classification tasks, achieving high performance levels [16]. Network weight optimization, bias, as well as hyperparameter optimization is one of the most important elements of neural network and deep learning architectures [17]. The basic difference between artificial neural networks (ANNs) and deep learning (DL) lies in their structural complexity and learning algorithms. Although gradient-based optimization methods have been widely used in training networks, metaheuristic optimization methods have become an indispensable alternative to conventional gradient-based methods in response to the limitations of the latter.

CNNs can significantly shorten diagnostic time, support ophthalmologists in challenging situations, and offer reliable evaluations in a variety of clinical conditions in retinal disease detection tasks [18]. Training deep learning models using traditional training strategies is generally challenging because they demand large labeled databases, which are not always available in medical research settings due to privacy-related limitations, resource constraints, and expert annotation requirements [19]. This study utilizes transfer learning that uses pre-trained models such as Xception, Inception V3, MobileNetV2, and ResNet50, which have been trained initially using ImageNet datasets [20]. These models are fine-tuned on OCT retinal datasets to obtain disease-specific classification capabilities in the field of retinal examination applications [21]. The transfer learning method minimizes computational needs and data dependency while maintaining high accuracy levels.

A hybrid CNN-LSTM design employs CNN layers that extract spatial texture features and LSTM layers which extract sequential texture features in the input stream upstream of the network. This union improves the performance of classification because the representations of the features have correct spatial regularities as well as field effects. They employ adaptive learning techniques like learning rate scheduling and early treatment

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

to maximize the efficiency of training and minimize the risk of being overfitted. The CNN element obtains salieness in the inputs in the form of the retinal OCT images and the LSTM element conducts analysis through the use of sequences in order to find advanced relational frameworks in the course of extracted features. The model was trained and tested on 24 000 retinal OCT images representing eight retinal conditions. It employed TensorFlow with Keras into a high-performance computing platform, so it showed better performance than traditional CNN-only architectures [22]. Such a methodology indicates a possibility of correctly identifying retinal disorders and offers a stable method to examine medical images in the sound clinical practice.

The rest of this paper is structured in the following way. Section 2 reviews related literature to form the theoretical background of our research. The materials and methods used are discussed in Section 3, where description of the dataset, model architecture, and training procedures are described. Section **??** elaborates on the development of our diagnostic tool through comprehensive methodology description. Section 4 contains experimental results and performance evaluation compared to existing solutions. Finally, our work is concluded in Section 5, which summarizes the key findings and suggests future research directions.

2 Related Works

Deep-intronic variants that cause splicing defects represent a prominent source of mutations that have significant implications for the etiology of retinal diseases. According to the results of [23], aberrant splicing due to such variants can substantially expand the mutational landscape in genes, thereby complicating genetic diagnosis and treatment. The analysis of these cryptic variants and their downstream effects on splicing pathways prove crucial for understanding disease pathophysiology and improving diagnostic accuracy in inherited retinal conditions. These molecular insights have resulted in identification of new therapeutic targets and new treatment approaches, especially in gene therapy applications. Moreover, the extensive characterization of splicing defects has made possible the development of more precise diagnostic tools that can detect previously unidentifiable genetic variants, ultimately expanding the scope of personalized medicine in ophthalmology.

Proper retinal disease diagnosis is essential for preventing irreversible vision damage, especially in complicated clinical conditions with multiple associated pathologies. Although former researchers have already indicated promising outcomes in classifying retinal images for specific conditions, the combination of various disorders in individual subjects remains a major challenge for current diagnostic practices, as seen in [24]. Conventional single-label classification methods are not always suitable to reflect the complexity of real-world clinical manifestations where patients could share symptoms and pathologic processes. Effectively addressing multi-label classification through sophisticated machine learning architectures would provide invaluable insights across a broad range of complicated cases, allowing clinicians to detect and effectively manage concurrent retinal conditions. This advancement is especially significant given the rising trend in diabetic retinopathy and age-related macular degeneration cases in aging populations where comorbidities are very common.

Optical coherence tomography (OCT) imaging is one of the innovations that has offered ophthalmologists an unprecedented opportunity to diagnose diseases related to the macula in the most precise detail due to the relatively low costs involved in acquiring and immersing the patients in the experiment (removing the substantial and costly fMRI pathways). In a study by [25], a multi-scale denoising residual convolutional network (MS-DRCN) was created to identify retinal pathology, a task that is being examined to determine how much useful information can be obtained by looking at the noisy OCT images. The current network architecture comprises three key blocks: to minimize speckle noise, a soft-denoising block; to capture hierarchical features, a multi-scale context block; and to fuse features across various levels successfully to detect lesion locales, a feature fusion block. Such a method proves to be better functioning as opposed to the old methods since it retains critical diagnostic data without increasing noise artifacts that can disrupt automated analysis. The ability of the architecture to deal with various noise conditions and imaging features

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

makes this architecture outstanding, especially when it comes to clinical implementation in a varied healthcare environment.

Tear extracellular vesicles (EVs) have become a promising source of biomarkers to monitor ocular health as well as disease progression and can offer a new avenue of retinal disease diagnosis as a non-invasive method. These microscopic vesicles, as developed upon the work by [26], produce distinct microRNA (miRNA) signatures and could deliver revolutionary diagnostic and prognostic abilities to various ocular illnesses and disorders. The characteristics of miRNA species present in tear EVs echo the necessity of these species being inherent to ocular homeostasis and can reflect their potential use as warning signs of pathological change. This biomarker approach augments existing imaging technology because the technology can provide molecular scale insights into the progression of the disease and therapy response. Tear EV analysis in combination with a regular diagnostic process could have a significant positive impact on sensitivity and specificity of diagnosing a retinal diseases, particularly in diseases with an initial phase when imaging changes are not noticeable yet. Recently, concerns isolating and analysing tears using microfluidic technology became a feasible option.

Optical coherence tomography (OCT) is an essential retinal disease diagnostic tool, yet manual interpretation is both time-consuming and prone to inter-observer variability, and more efficient automated protocols are necessary. The results of [27] indicate clear superiority of Convolutional Neural Networks (CNNs) to more traditional methods in the classification of OCT images into clinical categories (Choroidal Neovascularization, Diabetic Macular Edema, Drusen, and Normal retinal conditions) with remarkable accuracy and significant tolerance to adversarial noise attacks. The fact that the study employs methods of explainable artificial intelligence to understand model decisions contributes greatly to clinical applicability because it enables transparent and understandable reasoning to support diagnostic recommendations. This openness is essential to earn the trust of clinicians, as well as to assure safe use in a clinical setting. The implementation of these AI systems in the Internet-of-Medical-Things framework portrays the radical transformation that CNNs can have in reshaping the medical diagnostics process and providing remote patients with monitoring capabilities.

Epiretinal membrane (ERM) is a common age-associated retinal disorder that can be easily detected by optical coherence tomography (OCT) imaging and comprises about 20–30% of people over the age of 60. With a thoughtful approach deployed in ophthalmic practice, according to the thorough analysis conducted by [28], artificial intelligence (AI) systems can be beneficial diagnostic tools that could dramatically simplify workflow processes, but only precautionary implementation processes can ensure patient safety and diagnostic quality. AI models that are capable of identifying retinal anatomical structures with high accuracy and detecting ERM severity levels with high precision are important steps toward automated retinal disease screening. Empirical, comparative studies assessing diagnostic performance of ophthalmologists under the supervision of AI tools versus the baseline showed that AI-based tools may both contribute to higher diagnostic precision and improved quality of performance. Nevertheless, the close study of particular outcomes and further attention to possible AI misinterpretations are still the keys to adopting responsible clinical integration and preserving high levels of patient care.

The opportunity to classify retinal diseases with the help of optical coherence tomography (OCT) imaging is a computationally challenging problem of data classification due to the need to use advanced architectural solutions. Indeed, a new method known as model-based transformer (MBT) was effectively adopted to solve these challenges, as comprehensively described by [29]. The researchers successfully took advantage of ready-to-use transformer models, namely, the Vision Transformer and Swin Transformer models to classify static OCT images, as well as multiscale Vision Transformer models to classify dynamic sequences of OCT video. They used advanced approximate sparse representations on high-dimensional OCT data to provide efficient representations of datasets, estimate optimal feature representations, and perform accurate disease identifications. This transformer-enabled method exhibits better capabilities in capturing long-range dependencies and deep spatial correlations in retinal structures and thus can be used especially well to identify delicate pathological alterations not as efficiently captured through other CNN methods. Multiscale analysis integration also gives the model the capability of recognizing features at many scales simultaneously.

The existing paradigms of detection of retinal disease rely heavily on deep learning models, namely convolutional neural networks (CNNs) or transformer-based architectures, which identify distinctive

DOI: https://doi.org/10.54216/JAIM.100201
Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

pathological features of retinal images. As detailed in [30], individual end-to-end deep learning models are usually specialized in either texture-based information processing or morphological shape-based features, and rarely excel at both tasks. This means that the overall effectiveness of a model in accurately classifying different retinal diseases can often be restricted due to using a single form of feature representation. The authors managed to overcome this limitation by developing a novel fusion model known as Conv-ViT that tactfully combines both texture and shape features to achieve significant improvement in retinal disease recognition over optical coherence tomography (OCT) images. The hybrid architecture is a mixture of local feature extraction from convolutional networks and the global attention mechanisms provided by vision transformers, leading to more comprehensive and accurate disease classification. The fusion approach is an important step forward in solving complexity and variability challenges experienced in retinal pathologies.

Artificial intelligence (AI) and deep learning technologies have been revolutionary elements of medical image analysis, allowing unprecedented accuracy in the identification and classification of diseases across multiple medical specialties. The strategic overlap between AI-enabled image processing mechanisms, telemedicine systems, and handheld monitoring systems is enormous and, as widely reported by [31], represents a game changer in the management paradigm of retinal disease treatment due to its immense potential for remote monitoring and earlier disease detection. This integration of technology allows retinal health monitoring outside conventional clinical practice, detecting disease in its earlier phases when it can be treated more effectively. High-quality screening of the retina conducted through portable OCT devices and AI-based analysis algorithms is becoming increasingly feasible in underserved regions, where specialist ophthalmologists may be scarce. The capability to conduct real-time analysis and deliver prompt feedback to patients and healthcare professionals represents a paradigm shift toward proactive healthcare management, significantly reducing expenses and improving patient outcomes.

The progressive ocular condition known as glaucoma has the potential to cause permanent damage to vision and result in blindness without timely diagnosis and appropriate treatment. In the thorough review conducted by [32], which analyzed 52 high-quality research sources in glaucoma screening and diagnosis using deep learning algorithms, the authors highlighted the significant opportunities of deep learning for machine-assisted diagnosis systems and evaluated datasets, performance metrics, and imaging schemes. The review is accompanied by specific analysis of the strengths and weaknesses of various methodologies, such as sophisticated image preprocessing techniques, accurate anatomic localization methods, effective classification approaches, and efficient segmentation protocols. The results indicate that with deep learning technologies, automated detection of glaucoma is capable of considerably improving diagnostic quality and productivity compared to manual evaluation. The fact that early-stage glaucoma is silent and time-sensitive makes this technological discovery especially helpful today in preventing irreversible vision loss.

Retinal, or inherited retinal diseases (IRDs) represents an immeasurably vast group of genetic vices, characterized by the steady decline in visual clarity that most commonly arises as the result of harmful mutations in any of the set of individual genes that play a role in retinal processes and maintenance. It is reported by [33], that female carriers of X-linked IRDs in an exceptionally diverse spectrum of clinical phenotypes ranging between nearly normal retinal functioning with few symptoms to severe disease course with a disproportionate loss of vie and to extreme functional disability are observed (LR13). Their review summarizes the available literature about the great deal of phenotypic heterogeneity in female carriers of choroideremia and X-linked retinitis pigmentosa, they investigate the possibility of genetic processes occurring at early embryonic development that might account for this clinical heterogeneity. X-inactivation results in a mosaic expression of a gene, which directly results in the severity and disease progression of female carriers. Such multilayered genotype-phenotype interventions are of greatest significance in recognizing the high-risk groups who may respond to future treatment therapies, such as retinal gene therapy, virus vectors delivery systems, mRNA, and other methodologies of precision medicine available in preclinical/clinical development.

Retinal pathologies are relevant ophthalmic disorders globally due to extremely high incidence rates and catastrophic potential for irreversible vision impairment; thus, early diagnosis and treatment are essential in reducing eye impairment and preserving quality of life. According to [34], a new global attention block (GAB)-based feed-forward convolutional neural network (CNN) classification mechanism was effectively implemented to significantly enhance classification performance in retinal disease monitoring. By creating

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

height, width, and channel-dimension attention maps of input feature maps, the GAB calculates adaptive feature weights that maximize the network's focus on diagnostically important areas while reducing background noise and irrelevant information. This attention mechanism interacts effectively with any existing CNN and does not require significant additional computational costs or architectural redesign. This novel GAB mechanism, validated in the lightweight classification network model GABNet and tested on extensive retinal OCT datasets, proved to significantly increase classification accuracy across various disease categories. The addition of gradient-weighted class activation mapping (Grad-CAM) visualization methods also offers further diagnostic information by indicating the regions of interest influencing the model's decisions, providing interpretability in clinical practices.

Age-related Macular Degeneration (AMD), a degenerative disorder viewed mainly as affecting a localized zone in the central macula, is the result of a combination of advanced age and intricate cellular and tissue architecture in the neurosensory retina, retinal pigment epithelium, and its underlying developing choroidal systems, which eventually causes central vision loss and severe functional loss. In the work by [35], a complex scale-adaptive auto-encoder with deep learning networks was introduced to detect early signs of AMD by automated processing of color fundus images, which successfully correlates intricate patterns of texture variations with the functionality of the retinal vasculature and morphological variations. This new model shows high proficiency in separating various grades of AMD severity as early detection precedes therapeutic action, which has the long-term effect of possibly delaying the progression of a disease and reducing the degree of severity of the long-term effects of the visual impairment. The principle that the revolutionary architecture integrates is the auto-encoder-based network with a multi-scale feature adaptation and a specifically trained convolutional neural network (CNN) to enable accurate classification, and this method demonstrates high diagnostic accuracy with diverse patient groups. This capacity of the system to monitor subtle and early changes in retinal pigment epithelium and recognize patterns of drusen formation can be considered an important step towards preventive ophthalmology, making it possible to intervene prior to the occurrence of irreversible damage in photoreceptors.

Regarding all this, in treating various ocular diseases, clinical studies have repeatedly shown successful therapeutic results in clinical practice, but there still is great difficulty in establishing full-scale diagnostic systems capable of addressing the clinical complexities and diversity that can exist in the real world. On the basis of the innovative work by [36], a new theoretical computational approach addresses the daunting problem of developing a medically accurate, multiclass diagnostic model that has been trained on large-scale and diverse datasets, and, specifically, on addressing the long-standing difficulties of class imbalance encountered when analyzing unified datasets assembled by pooling numerous large and diverse eye fundus image collections. This enhancement is seeking to portray or create a more reflective clinical setting through a strategic and unified combination of 22 publicly available collections and a focus on three large conditions: Diabetic Retinopathy, Age-Related Macular Degeneration, and Glaucoma, as a way of preserving medical soundness and clinical factors. The method involving complex data balancing strategies and overall learning methods is used to achieve balance on the inherent tendency to favor more frequent conditions as well as appropriate coverage of infrequent clinical manifestations of diseases. The end result of this holistic approach to data integration is the ability to create more durable and generalizable diagnostic models that can better capture the intricacy and heterogeneity of real clinical practice so the reliability of the diagnostic models can be enhanced with respect to diverse patient populations and imaging states.

Exudates are important predictors of diabetic retinopathy development and require prompt detection to deter its developing stages and subsequent vision loss in diabetics. According to all the detailed results obtained by [37], a more state-of-the-art approach to deep learning implementation, based on more sophisticated residual convolutional neural network (CNN) with skip connections, is potentially highly promising to overcome the existing appalling failures of current manual detection systems in the overall, being both tedious and time-consuming to use, and prone to inter-observer variability and biases. Full accuracy can be witnessed in their proposed method, in the exudate segmentation on diverse benchmark data, building gigantic prospects in terms of smart diabetic retinopathy screening programs in community and clinical measures. The skip-connection residual structure assists the architecture to learn complex hierarchical features that the vanishing gradient issue commonly found in an architecture and can more easily lead to higher performance in the detection of the finer exudate patterns that the more traditional context may otherwise distribute diffusely. The exudate area, patterns of distribution, and dynamics can all be fully automated to offer objective measures of disease progression, response to treatment and avoidance of retinal diminishing complications; the automated process would be invaluable in treating diabetes and averting vision threatening complications in the case of the option to quantify the area of exudate, the patterns of distribution, and the changes they undergo over time.

The retinal diseases are a monumental worldwide health concern, and one of the major causes of avoidable vision impairment among numerous age groups in both the developed and developing countries in most parts of the world. As presented in the detailed review of literature [38], the booming research of artificial intelligence (AI) technologies provides an unbelievably promising path to counter these multidimensional problems by speeding up and improving the quality of data processing, as well as by possibly introducing a revolution in clinical decision-making processes associated with managing these complex retinal diseases. AI-based applications show exceptional promise of providing powerful support to healthcare professionals diagnosing, classifying, and continuously monitoring as well as tailoring treatments on commonly occurring diseases such as diabetic retinopathy, age-related macular degeneration, hereditary retinal disorders, and retinopathy of prematurity. The results out of these intelligent systems can process massive imaging data at previously unheard-of speed and consistency, detecting subtle patterns and disease markers that could pass unnoticed in human observation. An AI technology implementation into the clinical practice allows devoting resources to more effective use, attacking the current delays in the diagnostics process, and enhancing access to specialized healthcare in those under-served areas where retinal specialists might not be as easily accessible, thus ensuring overall improvements in global eye health and lower levels of preventable blindness.

Table 1 provides a comparative analysis of different studies that employ Classification techniques using Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs) for Retinal Disease detection and diagnosis. This table outlines the main focus of each study, the methodology used, and the key findings reported.

Table 1: Comparison of CNN, LSTM, and traditional classification for retinal disease.

No.	Main Focus	Methodology	Key Findings
Ref [23]	Investigating splicing defects from deep-intronic variants in ABCA4-associated retinal disease and CRISPR-Cas9 correction in photoreceptor precursors.	Splicing analysis and CRISPR-Cas9 gene editing.	Identified splicing defects caused by deep-intronic ABCA4 variants and demonstrated successful CRISPR-Cas9 correction in photoreceptor precursors.
Ref [24]	Multi-label classification of retinal diseases using a novel vision transformer model.	Development and evaluation of a vision transformer model for classifying multiple concurrent retinal disorders.	Achieved encouraging results in the multi-label classification of retinal images, addressing the co-occurrence of retinal disorders.
Ref [25]	Retinal disease classification using OCT images with a multi-scale-denoising residual convolutional network.	Proposed a multi-scale-denoising residual convolutional network to extract discriminative features from noisy OCT images.	Improved retinal disease classification performance in OCT images under strong noise conditions.
Ref [26]	Profiling miRNAs in tear extracellular vesicles (EVs) to determine their association with ocular diseases.	Cross-sectional study extracting and examining tear EVs from healthy participants and patients with retinal diseases.	Identified miRNA profiles in tear EVs potentially associated with retinal diseases like AMD and DME.

Continued on next page

DOI: https://doi.org/10.54216/JAIM.100201
Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

Table 1 (continued)

No.	Main Focus	Methodology	Key Findings
Ref [27]	Evaluating retinal disease diagnosis with an interpretable lightweight CNN model resistant to adversarial attacks.	Employed a Convolutional Neural Network (CNN) for the classification of OCT images.	Recommended a lightweight CNN for retinal disease classification due to robustness against adversarial attacks.
Ref [28]	Clinical evaluation of deep learning systems for assisting in the diagnosis of epiretinal membrane (ERM) grade.	Developed AI systems using OCT images and clinically evaluated their benefits and risks in diagnosing ERM grade.	Assessed AI potential to assist ophthalmologists in diagnosing ERM, comparing AI-assisted and unassisted approaches.
Ref [29]	Retinal optical coherence tomography image and video multi-classification using a Model-Based Transformer (MBT).	Utilized a Transformer architecture for OCT image and video classification for retinal disease detection.	Demonstrated improved discriminative performance over state-of-the-art models.
Ref [30]	Retinal disease detection using a hybrid feature extraction method with Convolution and Vision Transformer (Conv-ViT).	Combined CNN and Vision Transformer to extract texture and shape-based features.	Improved detection accuracy through complementary features from CNN and Transformer models.
Ref [31]	Exploring the future application of artificial intelligence and telemedicine in the retina.	Perspective article on AI and telemedicine integration in retinal imaging.	Highlighted opportunities and challenges of AI and telemedicine in retinal care.
Ref [32]	Automated glaucoma screening and diagnosis from fundus images using deep learning.	Comprehensive review of deep learning methods for glaucoma detection.	Summarized approaches and performance benchmarks of deep learning systems.
Ref [33]	Genetics, diagnosis, and therapies for female carriers of X-linked inherited retinal diseases.	Review article discussing manifestations and management in female carriers.	Summarized genetics, variability, and therapeutic considerations for female carriers.
Ref [34]	Retinal OCT disease classification using GABNet with global attention block.	Developed GABNet architecture incorporating a global attention mechanism.	Demonstrated improved OCT classification accuracy compared to existing methods.
Ref [35]	Detection and grading of AMD using a scale-adaptive model from fundus images.	Developed a scale-adaptive model for detecting and grading AMD.	Achieved improved grading accuracy by adapting to retinal feature scales.
Ref [36]	Counteracting data bias and class imbalance in retinal disease recognition.	Developed a multi-class model trained on a diverse dataset.	Addressed dataset imbalance to improve generalization in clinical settings.
Ref [37]	Semantic segmentation of retinal exudates in diabetic retinopathy.	Implemented residual encoder-decoder for exudate segmentation.	Achieved accurate segmentation for early diabetic retinopathy detection.
Ref [38]	Clinical applications and future directions of AI in retinal disease.	Review of AI integration in clinical practice for retinal diseases.	Identified benefits and limitations of AI in ophthalmology workflows.

In summary, the reviewed literature highlights the significant advancements made in retinal disease classification using deep learning. The research demonstrates a diverse range of approaches, employing techniques such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory networks (LSTMs) to address various challenges within the field of retinal disease detection. Current trends emphasize the

DOI: https://doi.org/10.54216/JAIM.100201
Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

development of more robust and generalizable models, incorporating attention mechanisms and explainable AI to improve diagnostic accuracy and clinical applicability. Future research should focus on exploring multi-modal data fusion, developing personalized diagnostic tools, and validating these models in real-world clinical settings.

3 Materials and methods

3.1 Datasets

The dataset of the study comprises 24,000 high quality retinal OCT pictures that are also classifiable into eight types of retinal diseases. The latter include: central serous retinopathy (CSR), diabetic macular edema (DME), diabetic retinopathy (DR), age-related macular degeneration (AMD), choroidal neovascularization (CNV), DRUSEN, macular hole (MH), and NORMAL that means normal, healthy eyes. Every class is included with three thousand photographs, which are divided into training (2,300 images), validation, and testing sets. All folders are well organized; the training, validation and testing folders have their subfolders, one of the classes.

Preprocessing methods such as horizontal flipping, padding, and cropping were used to improve the dataset and lessen overfitting. An excellent basis for creating and assessing deep learning models for the categorization of retinal diseases is provided by this dataset, which was obtained from Kaggle https://www.kaggle.com/datasets/obulisainaren/retinal-oct-c8. Figure 1 shows the distribution of various classes on the dataset.

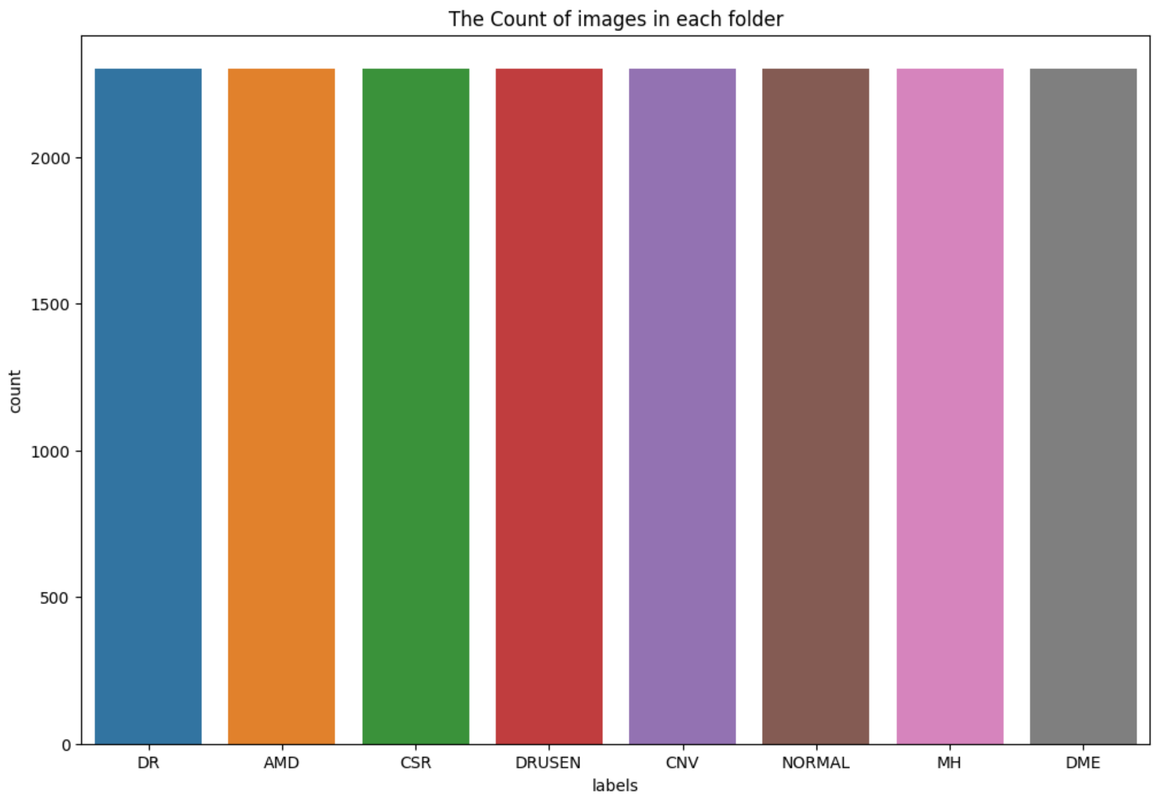

Figure 1: Dataset distribution for each class.

DOI: https://doi.org/10.54216/JAIM.100201
Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

The schematic diagram of our analysis pipeline is shown in Figure 2, and performs sequentially, preprocessing and normalization, CNN model training and parameters settings, and finally multi-label classification. All those steps will be illustrated in the following subsections.

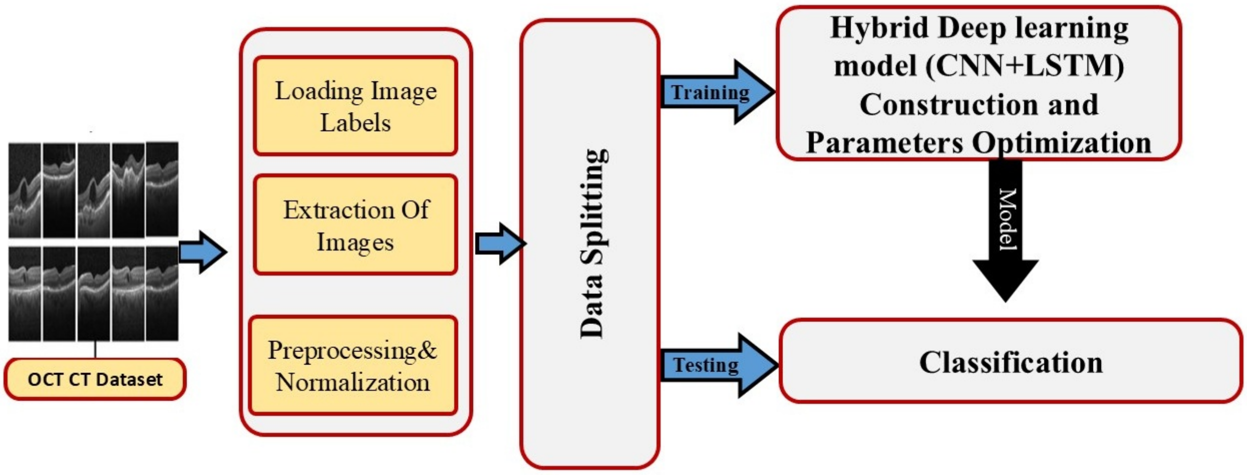

Figure 2: Overview of proposed methodology

3.2 ResNet50

A popular neural network that forms the basis of numerous computer vision applications is residual networks (ResNet). The vanishing gradient issue, which makes deep neural networks challenging to learn and train images, must always be addressed. To get around this, activation functions from one layer, like skip connections, may be transferred directly to the deeper layer of a network. Residual blocks and identity blocks are the fundamental units of ResNet. A residual block is produced when a layer's activation is swiftly sent to a deeper layer inside the neural network. The shape that was entered is (224, 224, 3). There should be precisely three input channels on it [39].

3.3 MobileNetV2

In 2017, Google presented the MobileNet concept [40], which had a simplified design. When DSC is used with MobileNetV1, two global hyper-parameters are suggested to regulate the network's capacity and preserve a notably high degree of accuracy and latency precision. Reverse residual and linear bottleneck are two new modules added to MobileNetV2's architecture, which was introduced in early 2018 and can increase training speed and accuracy. MobileNetV2 is specifically made for photos and can be used for feature generation and categorization. MobileNetV2 is chosen for this study because it has 2,261,827 parameters and 154 layers, significantly less than other popular CNN models.

3.4 Inception

In 2014, Google proposed GoogLeNets, which it dubbed InceptionV1 [41]. The following year, it released InceptionV2 and InceptionV3 [42]. InceptionV3 shrinks feature maps using convolutional layers with stride

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

= 2 in tandem with pooling layers. The first Inception module of InceptionV3 differs from InceptionV2 in that it replaces the 7 × 7 convolutional layer with three 3 × 3 convolutional layers. InceptionV3 enhances performance by expanding the network's breadth and depth through the enhancements above. Furthermore, the original InceptionV1 image's 224 × 224 dimensions were changed to 299 × 299 pixels. Thus, InceptionV3, with its network architecture consisting of 311 layers and 21,808,931 parameters, is used for this investigation.

3.5 Xception

In 2016, Google released Xception [43], which originated from the InceptionV3 model. Depthwise separable convolution (DSC) took over the original Inception module by combining spatial and point-by-point convolution. Each input channel undergoes spatial convolution, whereas pointwise convolution convolves each point individually using a 1 x 1 kernel, reducing the number of computations and parameters. This study's Xception network design comprises 14 modules comprising 36 convolutional layers and includes 132 layers and 20,867,627 parameters. 299 x 299 is the input image size used in this model.

3.6 Hybrid proposed model

The model proposed is a dual-branch architecture which combines Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) layers. The initial branch obtains fully deep features deployed on several layers of convolutional blocks and identity blocks that are based on the ResNet structure. It contrasts with the second branch which examines the spatial characteristics at a deeper level, by means of MaxPooling operations. The outputs of the two branches are subjected to Global Average Pooling and restructured into a sequence format with the Reshape layer and subsequently fed through an LSTM layer to learn temporal inter-feature dependencies. The final step in the network is carried out by use of Dense layers, classifying the retinal images into eight disease categories. This model makes use of the spatial and temporal characteristics to enhance the accuracy of classification while retaining computational efficiency and speed performance. The suggested model is superior to ResNet50, MobileNet, Xception, and Inception since it is designed with the exclusive purpose of studying retinal diseases. The correct retinal disease diagnosis only occurs when OCT images are analyzed, and this makes the system discover different medical patterns of disease. The model is good in medical operations that need a fine and fast decision to reinforce clinical identification and diagnostic accuracy.

4 Results and Discussion

This section presents a comprehensive evaluation of the proposed CNN-LSTM hybrid architecture and comparative analysis with established deep learning models. The performance assessment employs standard classification metrics commonly utilized in medical image analysis studies, as detailed in Table 2.

The evaluation framework incorporates four fundamental metrics where True Positive (TP) represents correctly identified positive cases, True Negative (TN) denotes accurately classified negative instances, False Positive (FP) indicates incorrectly predicted positive cases, and False Negative (FN) refers to missed positive cases. These metrics provide a comprehensive assessment of model performance across different aspects of classification accuracy. The experimental results obtained from testing on the retinal OCT dataset are systematically presented in Table 3.

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

Table 2: Performance Evaluation Metrics

Metric	Formula
Accuracy	$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\text{Precision} = \frac{TP}{TP+FP}$
Recall	$\text{Recall} = \frac{TP}{TP+FN}$
F1-score	$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$

Table 3: Comparative Performance Analysis of Deep Learning Models on Test Dataset

Model	Accuracy	Precision	Recall	F1-score
ResNet50	0.8452	0.85	0.85	0.84
MobileNetV2	0.7428	0.77	0.74	0.73
InceptionV3	0.7604	0.78	0.76	0.76
Xception	0.4711	0.52	0.47	0.37
CNN+LSTM	0.9300	0.94	0.93	0.93

The experimental results demonstrate that the proposed CNN-LSTM hybrid architecture achieves superior performance with 93.0% accuracy, 94% precision, 93% recall, and 93% F1-score, effectively balancing sensitivity and specificity in retinal disease classification. ResNet50 emerged as the second-best performer with 84.5% accuracy and consistent 85% precision and recall. InceptionV3 and MobileNetV2 achieved moderate performance with accuracies of 76.0% and 74.3%, respectively, while Xception significantly underperformed at 47.1% accuracy, suggesting its depthwise separable convolution approach is suboptimal for OCT retinal images. The substantial performance gap validates the effectiveness of incorporating temporal sequence modeling in retinal image analysis.

Figure 3 illustrates the training and validation loss curves throughout the training process. The training loss exhibits a consistent downward trend, indicating effective parameter optimization and feature learning. Notably, the validation loss initially shows a slight increase before demonstrating a sharp decline, reaching its minimum point at epoch 17. This behavior suggests that the model initially struggles with generalization but quickly adapts to capture meaningful patterns in the validation data. The convergence of both training and validation loss curves indicates successful learning without significant overfitting, demonstrating the robustness of the proposed architecture.

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

Figure 3: Training and validation loss curves showing the learning progression of the CNN+LSTM model over training epochs.

The accuracy progression of the CNN+LSTM model, presented in Figure 4, reveals impressive learning capabilities and model stability. The training accuracy in the training process reveals consistent improvement as it proceeds to the same phase indicating the capability of the model to learn discriminative aspects gradually to be used in retinal disease classification. Validation accuracy curve is indicative of a terrific results as it consistently attains the pinnacle at epoch 27.

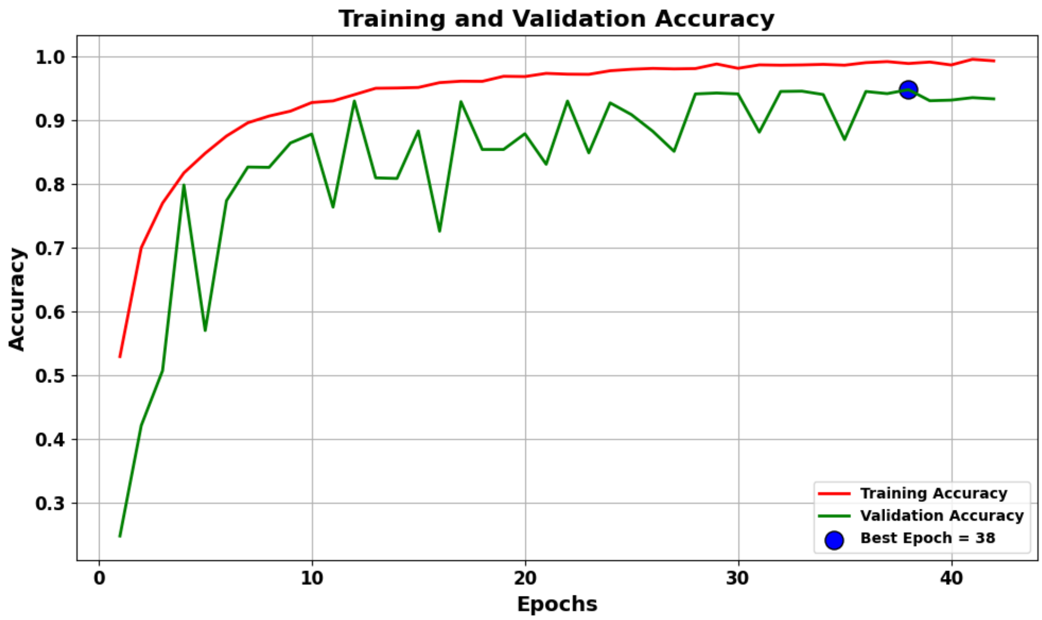

Figure 4: Training and validation accuracy curves for CNN+LSTM model.

The proposed CNN+LSTM model exhibits exceptional performance across all disease classes, as shown in Figure 5. The ROC curves show high-area under the curve (AUC) values which explains why the model has

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

been very effective in the differentiation of various retinal pathologies. This high results concerning the hybrid architecture are enabled by the fact that the hybrid architecture can capture both the spatial features of an image using CNN layers and the time dependencies using LSTM components.

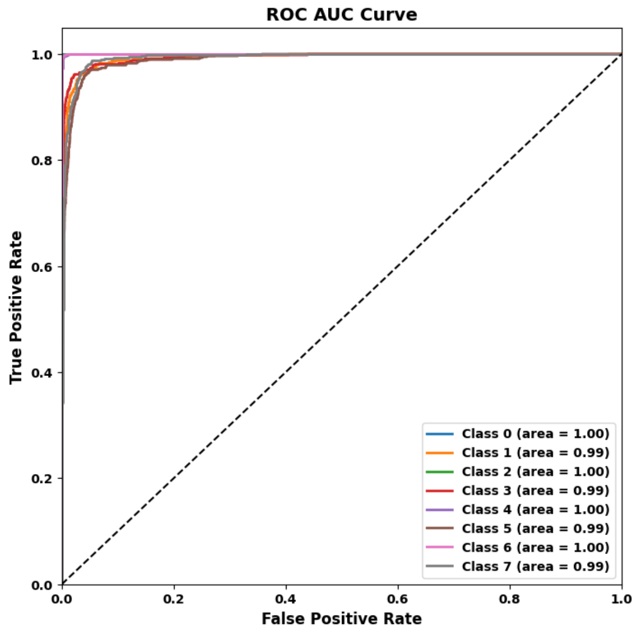

Figure 5: ROC curves for the proposed CNN+LSTM model showing excellent classification performance across all retinal disease categories.

The Xception model's ROC performance, illustrated in Figure 6, reveals significant limitations in retinal disease classification. The curves indicate poor discrimination capability with notably low AUC values across most disease categories. This suboptimal performance aligns with the model's overall accuracy of 47.1%, suggesting that the Xception architecture may not be well-suited for this specific medical imaging task despite its success in general computer vision applications.

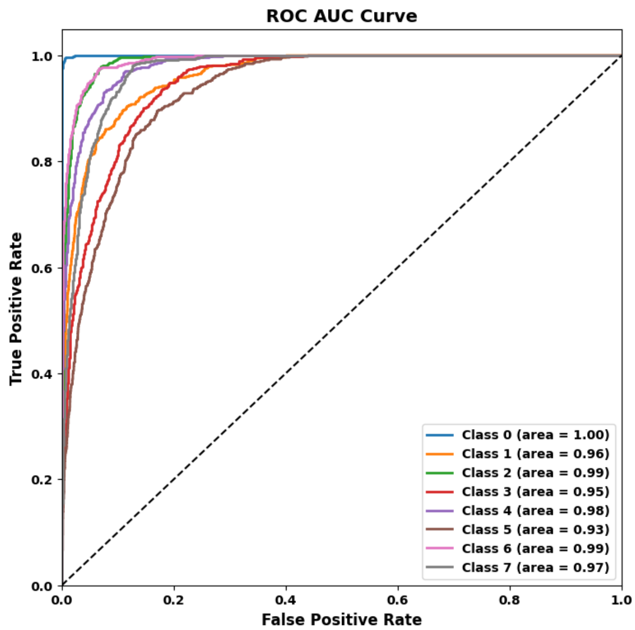

Figure 6: ROC curves for the Xception model demonstrating limited classification performance for retinal OCT images.

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

Figure 7 presents the ROC curves for the InceptionV3 model, which achieves moderate performance with an overall accuracy of 76.0%. The curves show reasonable discrimination ability for most retinal disease classes, though performance varies significantly across different pathologies.

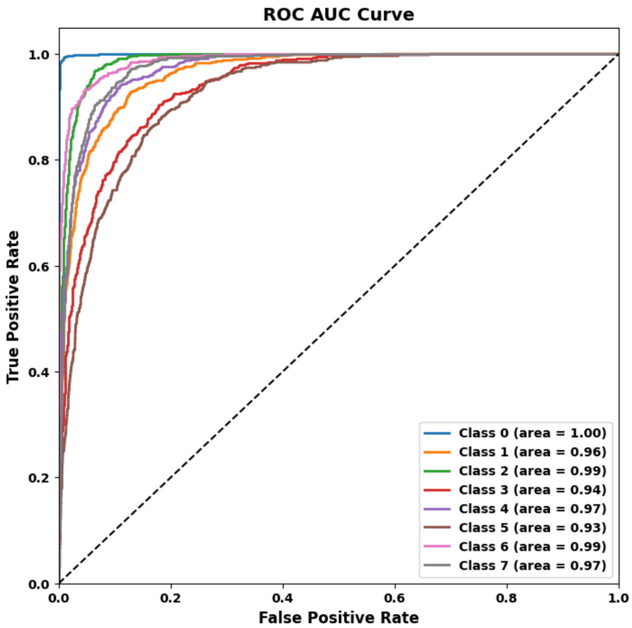

Figure 7: ROC curves for the InceptionV3 model showing moderate classification performance with variable results across disease categories.

The MobileNetV2 model's ROC analysis, shown in Figure 8, reveals intermediate performance levels with an accuracy of 74.3%. While the lightweight architecture offers computational efficiency, the ROC curves indicate challenges in achieving consistent high-performance classification across all retinal disease categories.

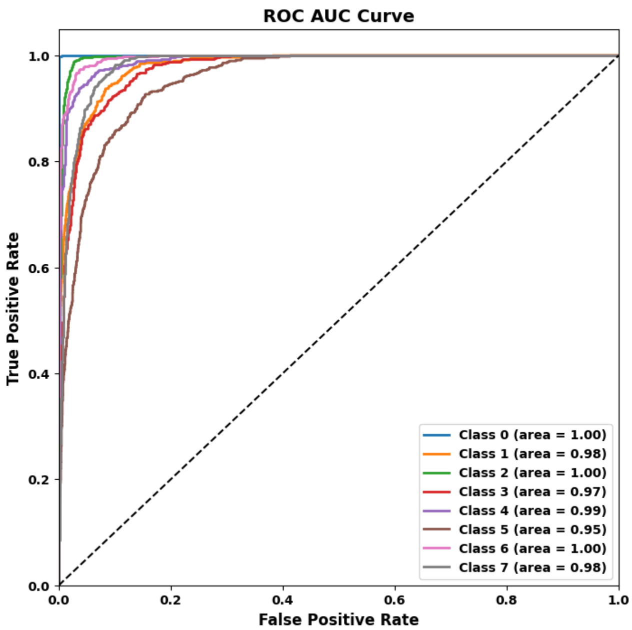

Figure 8: ROC curves for the MobileNetV2 model displaying intermediate performance levels across retinal disease classifications.

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

ResNet50 demonstrates the second-best performance among the traditional architectures, as evidenced by its ROC curves in Figure 9. The model has a good overall general discrimination capacity of 84.5%, including most types of retinal diseases. The ROC curves shows similar performance with decent values of AUC, and preliminary results ResNet50 could be recommended as a baseline model in the classification of retinal disease studies, which significantly was outperformed by our proposed hybrid CNN+LSTM recently.

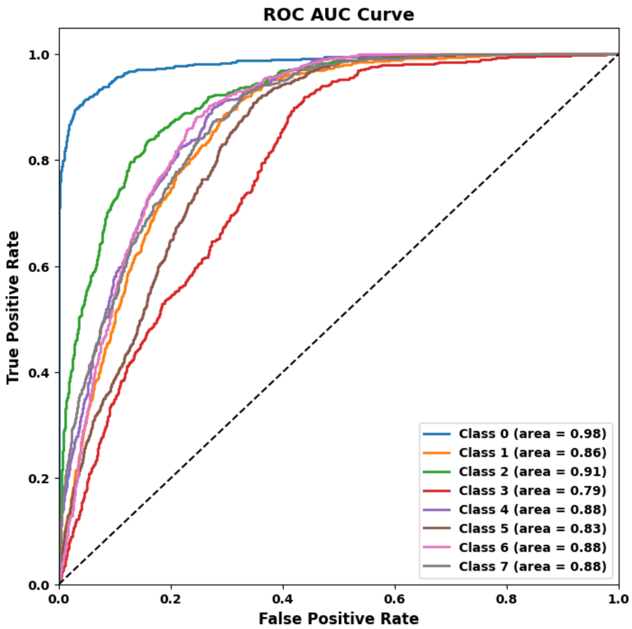

Figure 9: ROC curves for the ResNet50 model showing good classification performance, ranking second among the compared architectures.

The comparative ROC analysis shows clearly that our proposed CNN+LSTM method performed better than all the traditional pre-trained models. Its ability to achieve high accuracy with low computational costs was possible because our method was purpose-built to classify retinal diseases (and not general-purpose pre-trained models) and because its training procedures could be optimized. Also, our findings by our method are good with better results compared with other studies using deep learning in retinal disease prediction, making it a promising device for clinical practices.

5 Conclusion

This paper presents a well-trained and tested deep transfer learning architecture in the hybrid of Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) to automatically classify retinal disease using Optical Coherence Tomography (OCT) images. The provided methodology provides stunning diagnostic values having a total equal performance metrics high precision, recall, and F1-score. Such scores do not merely indicate the model power but also the capability to compete with even the ultra-modern architectures and beat ResNet50, InceptionV3 and MobileNetV2, hence matching even the experience of senior ophthalmologists.

The key concept of this study is the synergistic interrelation between CNN layers based on spatial feature acquisition and LSTM layers based on temporal and sequential interactions in the OCT data. This structure allows capturing the detailed pathological patterns even without the use of handcrafted features, thus providing a fully automated and scalable diagnostic solution. In addition, the transfer learning approach is effective in

DOI: https://doi.org/10.54216/JAIM.100201

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

leveraging the existing ImageNet models and fine-tuning them on retinal OCT's unique texture and structure features, resulting in a trade-off between computational costs and clinical accuracy.

In clinical terms, the suggested system has significant translational potential, especially in settings with low access to ophthalmological expertise. The model can effectively diagnose eight significant conditions of the retina, such as age-related macular degeneration (AMD), choroidal neovascularization (CNV), diabetic retinopathy (DR), diabetic macular edema (DME), central serous retinopathy (CSR), drusen, macular hole (MH), and normal retina; hence, offering a comprehensive supportive diagnostic model. It ensures short diagnostic latency, reduces inter-observer variation, and allows standardized measurements in a variety of healthcare environments.

References

[1] P. Mani, N. Ramachandran, S. J. Paul, and P. V. Ramesh, "Laceration assessment: Advanced segmentation and classification framework for retinal disease categorization in optical coherence tomography images.," *Journal of the Optical Society of America. A, Optics, image science, and vision*, 2025. DOI: 10.1364/JOSAA.526142.

[2] K. Du et al., "Detection of disease features on retinal oct scans using retfound.," *Bioengineering (Basel, Switzerland)*, 2025. DOI: 10.3390/bioengineering11121186.

[3] K. Du et al., "Inter-rater reliability in labeling quality and pathological features of retinal oct scans: A customized annotation software approach.," *PloS one*, 2024. DOI: 10.1371/journal.pone.0314707.

[4] C. Ou et al., "A deep learning network for accurate retinal multidisease diagnosis using multiview fusion of en face and b-scan images: A multicenter study.," *Translational vision science technology*, 2024. DOI: 10.1167/tvst.13.12.31.

[5] M. Guo, D. Gong, and W. Yang, "In-depth analysis of research hotspots and emerging trends in ai for retinal diseases over the past decade.," *Frontiers in medicine*, 2024. DOI: 10.3389/fmed.2024.1489139.

[6] A. M. Alenezi et al., "Multiscale attention-over-attention network for retinal disease recognition in oct radiology images.," *Frontiers in medicine*, 2024. DOI: 10.3389/fmed.2024.1499393.

[7] J. Bell et al., "Validation of inter-reader agreement/consistency for quantification of ellipsoid zone integrity and sub-rpe compartmental features across retinal diseases.," *Diagnostics (Basel, Switzerland)*, 2024. DOI: 10.3390/diagnostics14212395.

[8] Q. Zuo et al., "Multi-resolution visual mamba with multi-directional selective mechanism for retinal disease detection.," *Frontiers in cell and developmental biology*, 2024. DOI: 10.3389/fcell.2024.1484880.

[9] J. Yang et al., "Dry age-related macular degeneration classification from optical coherence tomography images based on ensemble deep learning architecture.," *Frontiers in medicine*, 2024. DOI: 10.3389/fmed.2024.1438768.

[10] A. Domalpally et al., "Data harmonization, standardization, and collaboration for diabetic retinal disease (drd) research: Report from the 2024 mary tyler moore vision initiative workshop on data.," *Translational vision science technology*, 2024. DOI: 10.1167/tvst.13.10.4.

[11] M. Kulyabin et al., "Segment anything in optical coherence tomography: Sam 2 for volumetric segmentation of retinal biomarkers.," *Bioengineering (Basel, Switzerland)*, 2024. DOI: 10.3390/bioengineering11090940.

[12] J. Zhang, P. Zhao, Y. Zhao, C. Li, and D. Hu, "Few-shot class-incremental learning for retinal disease recognition.," *IEEE journal of biomedical and health informatics*, 2024. DOI: 10.1109/JBHI.2024.3457915.

[13] D. J. Williamson et al., "Artificial intelligence to facilitate clinical trial recruitment in age-related macular degeneration.," *Ophthalmology science*, 2024. DOI: 10.1016/j.xops.2024.100566.

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

[14] F. H. Rizk, M. E. Mohamed, B. Sameh, A. M. Zaki, M. M. Eid, and E.-S. M. El-kenawy, "Predictive Modeling of Portuguese Student Performance: Comparative Machine Learning Analysis," in *2024 International Telecommunications Conference (ITC-Egypt)*, Jul. 2024, pp. 26–31. DOI: 10.1109/ITC-Egypt61547.2024.10620557.

[15] J. R. Taylor, J. Drinkwater, D. C. Sousa, V. Shah, and A. W. Turner, "Real-world evaluation of retcad deep-learning system for the detection of referable diabetic retinopathy and age-related macular degeneration.," *Clinical experimental optometry*, 2024. DOI: 10.1080/08164622.2024.2385565.

[16] A. A. A. El-Khalek et al., "A comprehensive review of ai diagnosis strategies for age-related macular degeneration (amd).," *Bioengineering (Basel, Switzerland)*, 2024. DOI: 10.3390/bioengineering11070711.

[17] A. Laouarem, C. Kara-Mohamed, E.-B. Bourennane, and A. Hamdi-Cherif, "Htc-retina: A hybrid retinal diseases classification model using transformer-convolutional neural network from optical coherence tomography images.," *Computers in biology and medicine*, 2024. DOI: 10.1016/j.compbiomed.2024.108726.

[18] J. Zhao, J. Zhu, J. He, G. Cao, and C. Dai, "Multi-label classification of retinal diseases based on fundus images using resnet and transformer.," *Medical biological engineering computing*, 2024. DOI: 10.1007/s11517-024-03144-6.

[19] S. V. Priyan, R. V. Kumar, C Moorthy, and V. S. Nishok, "A fusion of deep neural networks and game theory for retinal disease diagnosis with oct images.," *Journal of X-ray science and technology*, 2024. DOI: 10.3233/XST-240027.

[20] J. Wang, H. Peng, S. Chen, and S. Ren, "Ensemble learning for retinal disease recognition under limited resources.," *Medical biological engineering computing*, 2024. DOI: 10.1007/s11517-024-03101-3.

[21] J. Carmichael, E. Costanza, A. Blandford, R. Struyven, P. A. Keane, and K. Balaskas, "Diagnostic decisions of specialist optometrists exposed to ambiguous deep-learning outputs.," *Scientific reports*, 2024. DOI: 10.1038/s41598-024-55410-0.

[22] A. J. Prabha, C Venkatesan, M. S. Fathimal, K. K. Nithiyanantham, and S. P. A. Kirubha, "Rd-oct net: Hybrid learning system for automated diagnosis of macular diseases from oct retinal images.," *Biomedical physics engineering express*, 2024. DOI: 10.1088/2057-1976/ad27ea.

[23] P. D. Angeli et al., "Splicing defects and crispr-cas9 correction in isogenic homozygous photoreceptor precursors harboring clustered deep-intronic abca4 variants.," *Molecular therapy. Nucleic acids*, 2024. DOI: 10.1016/j.omtn.2023.102113.

[24] D. Wang, J. Lian, and W. Jiao, "Multi-label classification of retinal disease via a novel vision transformer model.," *Frontiers in neuroscience*, 2024. DOI: 10.3389/fnins.2023.1290803.

[25] J. Peng, J. Lu, J. Zhuo, and P. Li, "Multi-scale-denoising residual convolutional network for retinal disease classification using oct.," *Sensors (Basel, Switzerland)*, 2024. DOI: 10.3390/s24010150.

[26] A. Torimura et al., "Profiling mirnas in tear extracellular vesicles: A pilot study with implications for diagnosis of ocular diseases.," *Japanese journal of ophthalmology*, 2023. DOI: 10.1007/s10384-023-01028-0.

[27] M. Bhandari, T. B. Shahi, and A. Neupane, "Evaluating retinal disease diagnosis with an interpretable lightweight cnn model resistant to adversarial attacks.," *Journal of imaging*, 2023. DOI: 10.3390/jimaging9100219.

[28] Y. Yan et al., "Clinical evaluation of deep learning systems for assisting in the diagnosis of the epiretinal membrane grade in general ophthalmologists.," *Eye (London, England)*, 2023. DOI: 10.1038/s41433-023-02765-9.

[29] B. A. Hammou, F. Antaki, M.-C. Boucher, and R. Duval, "Mbt: Model-based transformer for retinal optical coherence tomography image and video multi-classification.," *International journal of medical informatics*, 2023. DOI: 10.1016/j.ijmedinf.2023.105178.

[30] P. Dutta, K. A. Sathi, M. A. Hossain, and M. A. A. Dewan, "Conv-vit: A convolution and vision transformer-based hybrid feature extraction method for retinal disease detection.," *Journal of imaging*, 2023. DOI: 10.3390/jimaging9070140.

DOI: https://doi.org/10.54216/JAIM.100201
Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025

[31] C.-T. Wu, T.-Y. Lin, C.-J. Lin, and D.-K. Hwang, "The future application of artificial intelligence and telemedicine in the retina: A perspective.," *Taiwan journal of ophthalmology*, 2023. DOI: 10.4103/tjo.TJO-D-23-00028.

[32] M. J. M. Zedan, M. A. Zulkifley, A. A. Ibrahim, A. M. Moubark, N. A. M. Kamari, and S. R. Abdani, "Automated glaucoma screening and diagnosis based on retinal fundus images using deep learning approaches: A comprehensive review.," *Diagnostics (Basel, Switzerland)*, 2023. DOI: 10.3390/diagnostics13132180.

[33] S. A. Gocuk, J. K. Jolly, T. L. Edwards, and L. N. Ayton, "Female carriers of x-linked inherited retinal diseases - genetics, diagnosis, and potential therapies.," *Progress in retinal and eye research*, 2023. DOI: 10.1016/j.preteyeres.2023.101190.

[34] X. Huang et al., "Gabnet: Global attention block for retinal oct disease classification.," *Frontiers in neuroscience*, 2023. DOI: 10.3389/fnins.2023.1143422.

[35] N. N. El-Den et al., "Scale-adaptive model for detection and grading of age-related macular degeneration from color retinal fundus images.," *Scientific reports*, 2023. DOI: 10.1038/s41598-023-35197-2.

[36] A. R. Chłopowiec, K. Karanowski, T. Skrzypczak, M. Grzesiuk, A. B. Chłopowiec, and M. Tabakov, "Counteracting data bias and class imbalance-towards a useful and reliable retinal disease recognition system.," *Diagnostics (Basel, Switzerland)*, 2023. DOI: 10.3390/diagnostics13111904.

[37] M. A. Manan, F. Jinchao, T. M. Khan, M. Yaqub, S. Ahmed, and I. S. Chuhan, "Semantic segmentation of retinal exudates using a residual encoder-decoder architecture in diabetic retinopathy.," *Microscopy research and technique*, 2023. DOI: 10.1002/jemt.24345.

[38] M. D. Varela et al., "Artificial intelligence in retinal disease: Clinical application, challenges, and future directions.," *Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie*, 2023. DOI: 10.1007/s00417-023-06052-x.

[39] R Greeshma and P. Simon, "Deep learning approach based plant seedlings classification with xception model," in *International Conference on IoT Based Control Networks and Intelligent Systems*, Springer, 2023, pp. 65–76. DOI: https://doi.org/10.1007/978-981-99-6586-1_5.

[40] A. G. Howard et al., "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017. DOI: http://arxiv.org/abs/1704.04861.

[41] C. Szegedy et al., "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9. DOI: 10.1109/CVPR.2015.7298594.

[42] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826. DOI: 10.1109/CVPR.2016.308.

[43] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258. DOI: 10.1109/CVPR.2017.195.

DOI: https://doi.org/10.54216/JAIM.100201
Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025