Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.18576/amis/190515

A Deep Learning-Based Eye Disease Diagnosis Using OCT Imaging and SE-Enhanced CNNs

Amani A. Slamaa

Information System Department, Faculty of Computers and Information, Luxor University, Luxor, 85951 Egypt

Received: 7 Feb. 2025, Revised: 21 May 2025, Accepted: 23 Jul. 2025
Published online: 1 Sep. 2025

Abstract: Vision loss remains a major global health concern, with cataract, glaucoma, diabetic retinopathy (DR) being leading, yet often symptomless, causes of preventable blindness, highlighting the urgent need for early, accessible, and cost-effective diagnostic solutions. This study introduces DEEPSIGHT, a deep learning-powered diagnostic system designed to automatically detect these three diseases using optical coherence tomography (OCT) imaging. The system aims to deliver both high diagnostic accuracy and practical usability in clinical settings. At its core is custom convolutional neural network (CNN) architecture, enhanced with attention mechanisms such as squeeze-and-excitation (SE) blocks to improve feature extraction. The model was trained on a diverse dataset of OCT images collected from public sources and clinical partners. Preprocessing steps—including normalization, contrast enhancement and data augmentation—were applied to improve robustness and reduce over fitting. A stratified 5-fold cross-validation strategy was used during training, with categorical cross-entropy loss and the Adam optimizer. DEEPSIGHT achieved over 94% accuracy, with precision and F1-scores exceeding 92% across all classes. To support clinical interpretability, Grad-CAM and saliency maps were integrated, allowing visualization of the image regions influencing model predictions. The system was deployed in a prototype diagnostic platform and validated on an independent clinical dataset, confirming its reliability and real-world applicability. While currently limited to three diseases and local deployment, future work will focus on cloud integration, broader diagnostic coverage, and real-time teleophthalmology support to enhance accessibility and scalability. This research contributes to the growing field of AI in healthcare and underscores the transformative potential of deep learning in vision science.

Keywords: Cataracts, Convolutional Neural Networks (CNN), Deep Learning, Diabetic Retinopathy, Glaucoma

1 Introduction

Vision loss is a significant global public health issue, affecting individuals across all age groups and socioeconomic backgrounds. Among the leading causes of visual impairment and blindness are cataract, glaucoma, diabetic retinopathy, and diabetic retinopathy (DR)—three conditions that, if detected early, can often be treated or managed to prevent permanent vision loss [1] and [2].However, these diseases are particularly insidious because they frequently progress without noticeable symptoms until reaching advanced stages. This makes early diagnosis not only critical but also a key determinant in preserving vision and improving treatment outcomes.

Diagnosing eye diseases is very difficult because its symptoms do not appear until the late stages. Cataract, glaucoma, diabetic retinopathy, and diabetic retinopathy cause damage the optic nerve and retina that mean ailment of the main eye nerve for sight. By 2030, the numbers of affected people by these diseases will duplicate from the current 170 million to an estimated 367 million. Diabetes with high blood sugar levels often damages the capillaries blood vessels which called Diabetic retinopathy by rate 80% of the causes of vision loss. Where Diabetes is the most common reason of these diseases and it is expected that increasing approximately 37.3 million to 56.3 million among patients aged 20-79 years with diabetes by 2040. Thereby, the incidence rate of these diseases increases and multiplying the resulting damages, the most dangerous of which is blindness, especially in developing countries or un-urban areas. Cataract is a clouding of the lens in the eye and cause blurred vision. In cataract surgery, the cloudy lens replaces with a human-made lens. Glaucoma is a group of conditions that can damage the optic nerve. Because of normal proteins that build up over time of people in the

* Corresponding author e-mail: amani.slamaa@fci.luxor.edu.eg

United States, there are more than 90% of American people have cataract. While in Egypt total rate of effected people with eye diseases was 4.49%. Cataract is responsible for around 51% of blindness worldwide according to World Health Organization. It affects more than 2.7 million people aged 40 or older in United States, 2% in the urban area and 9% in rural areas in Egypt and 11.2 million in India. Current projections estimate 80 million people around the world suffer from glaucoma, and it is expected to increase to more than 111 million by 2040 [19].

According to the World Health Organization, more than 2.2 billion people worldwide suffer from some form of vision impairment, with at least 1 billion cases being preventable or left unaddressed. Cataracts remain the leading cause of blindness globally, accounting for approximately 51% of all cases. Glaucoma—often referred to as the "silent thief of sight"—causes irreversible damage to the optic nerve and currently affects over 80 million people worldwide. Diabetic retinopathy, a microvascular complication of diabetes, impacts more than 80% of diabetic patients after two decades of disease progression and is the primary cause of vision loss among working-age adults. By 2040, DR is projected to affect over 56 million individuals globally [3] and [4].

Diagnosing these conditions poses a clinical challenge, not only due to the complexity of interpreting retinal images but also because of limited access to ophthalmologists, particularly in underserved regions such as the Middle East and North Africa [5]. While traditional diagnostic methods are effective, they are often constrained by high costs, limited availability, and the need for specialized expertise.

Optical coherence tomography (OCT) has emerged as a powerful, non-invasive imaging technique that provides high-resolution cross-sectional views of the retina. In recent years, the convergence of OCT with artificial intelligence (AI) has opened new possibilities for scalable, automated, and accurate diagnostic solutions. Deep learning (DL), and specifically convolutional neural networks (CNNs), has shown exceptional performance in image-based classification tasks, including medical diagnostics. CNN is a class of artificial neural network that has become dominant in various vision tasks. CNN interests across a variety of domains including radiology. CNNs can learn complex visual patterns and detect subtle pathological features in retinal images that may be missed by the human eye. CNN is neural network designed for image recognition, classification, and feature extraction. It consists of Input Layer which represents an image as a tensor of shape (height, width, channels) for a gray scale image in 28×28×1, or a color image in 32×32×3, then convolutional Layer(s) that applies learnable filters (kernels) to the input to extract features like edges, textures, and patterns. Each filter slides over the image via stride producing a feature map. This layer followed by Activation Function (like ReLU) and Batch Normalization (optional). Pooling Layer(s) like MaxPooling is the next layer in CNN, it reduces spatial dimensions -width and height- and makes computation efficient and controlling over fitting. Flatten layer transforms feature maps into a vector, followed by Fully Connected (Dense) Layer(s) is one or more dense layers perform high-level reasoning and classification, it often ending in Softmax (for multi-class classification) or Sigmoid (for binary classification). Final layer of CNN is output Layer which produces the final predictions of image's class [22] Figure (1) illustrates the typical CNN model. This paper introduces DEEPSIGHT, a novel AI-powered diagnostic system that combines CNNs with OCT imaging to detect and differentiate between cataracts, glaucoma, and diabetic retinopathy. The system achieves a reported diagnostic accuracy of 93%. Designed as a user-friendly application, DEEPSIGHT facilitates seamless collaboration between physicians and radiologists, streamlining image acquisition, disease classification, and report generation. By addressing the limitations of conventional diagnostic tools and leveraging the capabilities of deep learning, DEEPSIGHT aims to empower healthcare providers with an efficient and accessible solution for early detection of vision-threatening diseases—ultimately contributing to better patient outcomes and a reduction in global blindness.

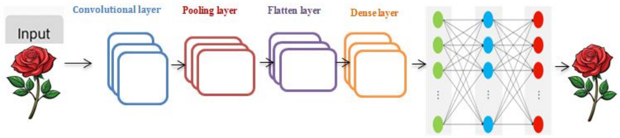

Fig. 1: CNN architecture

The contributions of this research are:

–Development of a CNN-based diagnostic model achieving 96% accuracy and precision, recall, and F1-scores all exceeding 92% .in classifying three major eye diseases; cataracts, glaucoma, and diabetic retinopathy. These results were validated through cross-validation and tested on an independent clinical dataset.

–The system was evaluated using real patient data from a private ophthalmic center. This helped confirm that DEEPSIGHT performs reliably across different imaging devices and patient populations.

–DEEPSIGHT is a custom convolutional neural network that incorporates attention mechanisms like Squeeze-and-Excitation blocks. These additions help the model focus on the most relevant parts of the image, improving its ability to detect subtle disease patterns.

- Integration of OCT imaging with AI to enhance diagnostic precision and reduce reliance on manual interpretation to meet clinical needs.
- To help clinicians understand how the model makes decisions, DEEPSIGHT includes visual explanation tools such as Grad-CAM and saliency maps. These tools highlight the image regions that influenced the model's predictions, making the system more transparent and easier to trust.
- Support ophthalmologists and general practitioners in early disease detection, ultimately contributing to improve patient outcomes and reduced global blindness burden.

The novelty of this paper is as follows:

- DEEPSIGHT can identify three major eye diseases from a single OCT scan. This makes it a more versatile and efficient solution for eye care.
- Many AI models are only tested in lab settings. DEEPSIGHT stands out by being evaluated in a real clinical workflow, which adds credibility to its practical value.
- By integrating visual explanation tools directly into the diagnostic process, DEEPSIGHT addresses one of the biggest concerns in AI healthcare: the lack of transparency. This helps bridge the gap between machine learning and clinical decision-making.

The paper outlines in section 2 displays background and literature review, section 3 proposes the methodology while the CNN model and its result addresses in subsections 3.1 and 3.2. The case validation, statistical analysis and discussion are handled in sub-sections 3.3 and 3.4. Finally conclusion is addressed in section 4.

Statement of Significance table

Problem or Issue	Vision-threatening diseases like cataract, glaucoma, and diabetic retinopathy often progress silently, making early diagnosis difficult and costly.
What is Already Known	Deep learning models have shown promise in medical imaging, but many are limited to single diseases, lack clinical validation, or require complex infrastructure.
What this Paper Adds	This study introduces DEEPSIGHT, a lightweight, explainable CNN-based system that detects three major eye diseases from OCT images with over 96% accuracy. It is validated on real clinical data and designed for practical use.
Who Would Benefit	Ophthalmologists, general practitioners, radiologists, and healthcare providers in underserved areas will benefit from this accessible, accurate diagnostic tool.

2 Background and Literature Review

The integration of artificial intelligence (AI) into ophthalmology has significantly advanced the early detection and diagnosis of retinal diseases. Numerous studies have explored the application of deep learning, particularly convolutional neural networks (CNNs), in analyzing retinal images obtained through modalities such as fundus photography and optical coherence tomography (OCT).

One such study demonstrated the effectiveness of CNNs and hybrid models in diagnosing conditions like diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma. Wang et al. proposed a hybrid deep learning framework that combines fully supervised and semi-supervised reciprocal learning to enhance both the accuracy and interpretability of OCT-based diagnosis. Their model achieved state-of-the-art performance in detecting AMD and diabetic macular edema (DME), aided by attention mechanisms and [6]. Fang et al. presented a lesion-aware CNN that enhances classification performance by focusing on lesion regions in OCT images. The model incorporates a lesion attention module that guides the network to learn more discriminative features, significantly improving classification accuracy for retinal diseases such as AMD and DME [7].In a related study, they also developed a method combining deep learning and graph search to segment nine retinal layers in OCT images of non-exudative AMD patients, enabling more precise structural analysis. It achieved 84.31% accuracy but is constrained by limited datasets and reduced effectiveness in early-stage detection [8].

The Yanbao Mobile Application is designed for glaucoma detection using SVM, random forest, and GBDT classifiers. It is portable and user-friendly but limited to a single disease and achieves only 77.31% accuracy. It also requires specialized hardware, which limits its scalability [14]. Yu and Dong presented RetinaDNet, an ensemble-based system that integrates fundus and vascular structure images. Their model achieved 99.2% accuracy in diabetic retinopathy detection and 98.8% in general retinal disease classification, demonstrating the power of dual-branch architectures and transfer learning [9].

Other research presents a lightweight and efficient deep learning model based on EfficientNetB0 for classifying ocular diseases from fundus photographs. The model was optimized through advanced preprocessing techniques and deployed on a web-based platform to improve accessibility. It achieved superior performance in diagnosing conditions such as myopia, hyperopia, astigmatism, glaucoma, and diabetic retinopathy. The integration into a user-friendly online system highlights its potential for real-time, remote ophthalmic diagnostics. Furthermore, another study in Neural Computing and Applications proposed a hybrid CNN-LSTM model for classifying dry and wet AMD using OCT images. By capturing both spatial and temporal features, the model improved early detection accuracy and demonstrated the potential of sequential modeling in ophthalmic diagnostics [10].

Chetoui and Akhloufi developed a CNN-based model for classifying OCT images into four categories: choroid neovascularization, diabetic macular edema, drusen, and normal. The model achieved an accuracy of 98.46% and an AUC of 0.998. It also incorporated explain ability features such as saliency maps to highlight lesion areas, enhancing clinical interpretability [11]. This comprehensive review [12] outlines the landscape of machine learning applications in ophthalmology, emphasizing the growing role of OCT imaging and CNN architectures. It highlights both the potential and the limitations of current AI-based diagnostic systems. The Diagnose-Me System utilizes deep learning models such as VGG16, ResNet50, and Xception to analyze smartphone-acquired fundus images. While it achieves 86% accuracy, it is limited by hardware dependency, computational inefficiency, and training complexity [13]. Adamopoulou et al. developed model for diagnosing eye diseases. The paper highlights how these technologies are being used to analyze complex ocular images with greater speed and accuracy than traditional methods. By applying models like convolutional neural networks (CNNs), the researchers demonstrate how AI can detect subtle structural changes in the retina and optic nerve, which are crucial for identifying conditions such as cataracts, glaucoma, and diabetic retinopathy. The study emphasizes that these tools not only support ophthalmologists in making more accurate diagnoses but also help bridge the gap in care for patients in underserved areas. The authors advocate for integrating AI into clinical workflows to improve diagnostic efficiency and accessibility, and they stress the importance of keeping pace with technological advancements to ensure better outcomes for patients [17]. G. ARSLAN et al. proposed CNN with 10-fold cross-validation to assess robustness for multi-class classification of three eye diseases; Cataract, Diabetic Retinopathy, Glaucoma and healthy cases with dataset contains 2748 retinal fundus photos (1374 normal and another 1374 images are spitted for three classes. The study compared five CNN architectures: DenseNet, EfficientNet, Xception, VGG, and ResNet. The proposed model performed with 91% for accuracy while the others get 85.35%, 94.8%, 68.28%, 83.08%, and 65.87 in order [21].

Despite the promising advancements in deep learning for ophthalmic diagnostics, several limitations persist across the reviewed studies. A common challenge is the limited diversity and size of datasets. Many models are trained on data collected from a single institution or imaging device, which restricts their generalizability to broader populations or different clinical settings. This lack of external validation raises concerns about over fitting and the robustness of these models when applied in real-world scenarios. Additionally, some studies rely on relatively small sample sizes, particularly for rare diseases or advanced imaging modalities like OCT angiography, which further limits the statistical power and reliability of their findings.

Technical and computational constraints also pose significant barriers. Deep learning models, especially those employing ensemble or hybrid architectures, often require substantial computational resources for training and deployment. Moreover, while OCT provides rich volumetric data, many models simplify this by using 2D slices, potentially overlooking critical spatial information inherent in 3D scans. This simplification can reduce diagnostic accuracy, particularly for diseases that manifest in subtle structural changes across retinal layers.

Another major limitation is the lack of standardization in imaging protocols and evaluation metrics. Variability in OCT acquisition settings, image resolution, and preprocessing techniques across different devices complicates model training and reproducibility. Furthermore, inconsistent reporting of performance metrics—such as accuracy, sensitivity, specificity, and AUC—makes it difficult to compare results across studies or benchmark progress in the field.

From a clinical perspective, interpretability remains a key concern. Although some models incorporate attention mechanisms or saliency maps to highlight relevant features, many still function as ”black boxes,” offering limited transparency into their decision-making processes. This lack of interpretability can hinder clinician trust and slow adoption in clinical practice. Additionally, most systems remain at the proof-of-concept stage and have not been integrated into clinical workflows or subjected to regulatory evaluation, limiting their immediate impact on patient care.

Finally, issues of generalization and bias are prevalent. Models trained on images from specific OCT machines may not perform well on images from other devices due to differences in imaging characteristics. Similarly, the underrepresentation of certain demographic groups in training datasets can lead to biased predictions, potentially exacerbating health disparities.

These systems highlight the potential of AI in ophthalmology but also reveal gaps in accuracy, accessibility, and disease coverage. These studies underscore the potential of AI in ophthalmic diagnostics but also highlight critical limitations in disease coverage, diagnostic accuracy, and system usability. The need for a comprehensive, accurate, and accessible diagnostic tool remains unmet—motivating the development of DEEPSIGHT.

In contrast, the proposed system—DEEPSIGHT—aims to address these limitations by combining a lightweight, efficient CNN architecture with robust OCT imaging and a dual-interface platform for both clinicians and radiologists. Moreover, unlike earlier models, DEEPSIGHT has undergone clinical validation on a real patient dataset, allowing for both technical evaluation and clinical relevance assessment.

3 Materials and Methodology

In this study, it introduces a deep learning-based approach for the automated detection of three major eye diseases—cataract, glaucoma, and diabetic retinopathy (DR)—using optical coherence tomography (OCT) imaging. The goal is to develop a system that is not only accurate and efficient but also practical for clinical use.

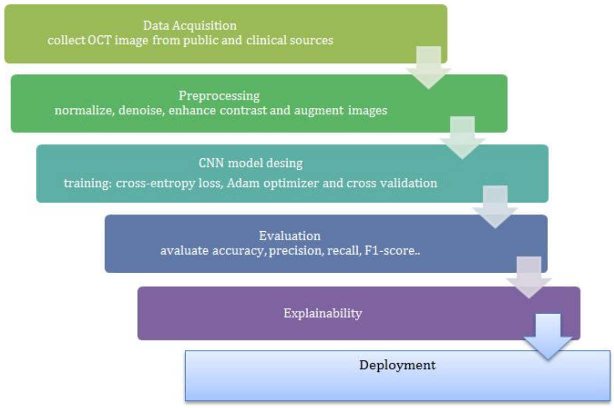

Fig. 2: main component of proposed methodology for DEEPSIGHT Application

The next figure (2) visualizes the main idea of this methodology. To begin, firstly compiling a diverse dataset of OCT images sourced from both public databases and clinical collaborators. This dataset will include labeled images representing each of the three target conditions, as well as healthy controls. To ensure consistency and improve model performance, all images will undergo preprocessing steps such as normalization, noise reduction, and contrast enhancement. That will also apply data augmentation techniques like rotation, flipping, and zooming to increase variability and reduce the risk of over fitting. The core of proposed system is a custom convolutional neural network (CNN) specifically designed for analyzing OCT images. The architecture will include multiple convolutional layers with ReLU activation and batch normalization, along with residual connections to maintain gradient flow during training. To help the model focus on the most relevant features, supported with attention mechanisms such as squeeze-and-excitation (SE) blocks or convolutional block attention modules (CBAM). A global average pooling layer will be used to reduce the feature map dimensions while preserving spatial information, followed by fully connected layers for final classification into one of four categories: cataract, glaucoma, DR, or normal. To make DEEPSIGHT's predictions easier to understand, visual explanation tools like Grad-CAM and saliency maps are added. These tools help show which parts of the OCT images the model is paying attention to when making a diagnosis.

Grad-CAM creates heat maps that highlight the most influential areas in an image, giving doctors a way to check if the model is focusing on the right features. Saliency maps go a step further by showing how sensitive the model's output is to each pixel, offering a detailed view of what's driving its decisions.

For training, the system will use a stratified 5-fold cross-validation strategy to ensure the model generalizes well across different subsets of the data. The training process will use categorical cross-entropy as the loss function and the Adam optimizer with a learning rate scheduler. The plan was train the model for up to 100 epochs with a batch size of 32, using early stopping to prevent over fitting. To address class imbalance, focal loss and class weighting where necessary will apply.

To evaluate the model's performance, the paper will use a comprehensive set of metrics including accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC) and Mathew's coefficient correlation (MCC) as used in [15], [16] and [20]. To make the system more transparent and clinically useful, it will integrate explain-ability tools such as Grad-CAM and saliency maps, which highlight the areas of the image that most influenced the model's decision. These visualizations will be reviewed by ophthalmologists to ensure they align with clinical understanding.

Finally, the proposed system deploys the trained model in a prototype diagnostic platform with user interface for clinicians to support real-time diagnosis and visualization, and to monitor performance and gather feedback. To validate the system's real-world applicability, it will tested on an independent clinical dataset and also assess the system's computational efficiency and compatibility with standard ophthalmic imaging hardware to ensure it can be realistically implemented in clinical environments.

Dataset Description

The Eye Diseases Classification dataset [18], sourced from Kaggle, comprises a comprehensive collection of over 4,000 high-resolution retinal fundus images, systematically categorized into four diagnostic classes: cataract, glaucoma, diabetic retinopathy, and Normal. Each image is stored in JPEG format and exhibits variability in resolution and quality, reflecting the heterogeneity commonly encountered in clinical imaging environments. This dataset has been widely adopted in machine learning research, particularly in the development of deep learning models for multi-class classification tasks. Prior studies have leveraged convolutional neural networks (CNNs) and transfer learning techniques to achieve high diagnostic accuracy using this dataset. To prepare the data for model training, standard preprocessing steps—such as image resizing, normalization, and data augmentation (e.g., rotation, flipping, and brightness adjustment)—are typically applied to enhance model robustness and generalization. Evaluation metrics commonly reported include accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC), with recent implementations achieving classification accuracies exceeding 90%. Given its diversity and clinical relevance, this dataset serves as a valuable benchmark for advancing AI-based diagnostic systems in ophthalmology.

Eligibility criteria of dataset

There are many factors for selecting dataset particularly for clinical images in healthcare field such as

Modality-Specific Imaging, Labeled Clinical Categories, Volume and Diversity, image Quality and Consistency, Public Accessibility and Reproducibility and finally, Relevance to Clinical Practice and AI Research. These features are applied and The Eye Diseases Classification dataset [18] is selected for the following reasons:

1. The dataset consists of high-resolution OCT images and are captured cross-sectional views of retinal layers.
2. The dataset is well-structured where it includes annotated classes for multiple common eye conditions allowing for multi-class classification aligned with diagnostic tasks.
3. With available size of samples per class, the dataset supports robust CNN training and minimizes risks of over fitting. The diversity of cases and conditions aids in well training models with trusted diagnostic performance.
4. As the dataset is free available and publicly hosted on Kaggle, it facilitates reproducibility, transparency, and providing benchmarking against other OCT-based diagnostic models.
5. The represented diseases are among the most prevalent causes of visual impairment globally, that enables massiveness of their images in dataset and applied for clinical AI research in ophthalmology.

Data Splitting Strategy

To ensure rigorous model evaluation and mitigate the risk of over fitting, the dataset was partitioned into three subsets: training, validation, and testing. Specifically, 70% of the images were allocated to the training set to facilitate model learning, 15% were designated for validation to support hyper parameter tuning and monitor performance during training, and the remaining 15% were reserved as an independent test set to assess the model's generalization capability. The splitting process was conducted using a stratified sampling approach to preserve the original class distribution across all subsets, thereby maintaining balance and reducing potential bias. This methodology ensures that each disease category is proportionally represented throughout the model development pipeline, supporting fair and consistent evaluation across all diagnostic classes.

3.1 Proposed System

The proposed DEEPSIGHT system supports diagnosing cataract, glaucoma, diabetic retinopathy through automated analysis, illustrated in figure (3), includes: The designed convolutional neural network (CNN) is tailored specifically for analyzing OCT images to detect common eye diseases—namely cataract, glaucoma, and diabetic retinopathy—alongside healthy cases. The model starts with a simple input layer that takes in gray scale images sized at 128 by 128 pixels, which is a practical resolution for balancing detail and computational efficiency.

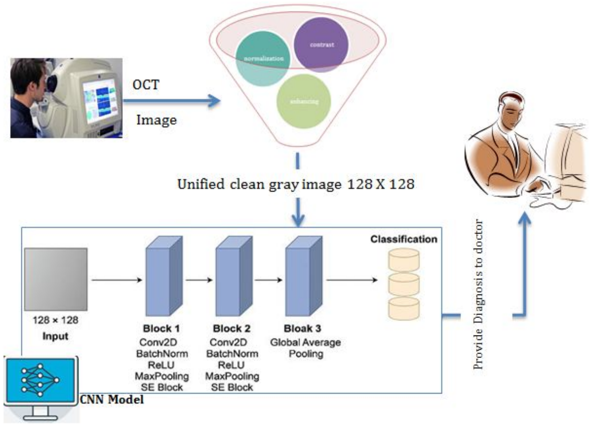

Fig. 3: proposed DEEPSIGHT system component

The network is built in three main stages, each designed to progressively extract more complex features from the images. In the first stage, the model uses a convolutional layer with 32 filters to detect basic patterns, followed by batch normalization and a ReLU activation to stabilize and speed up learning. A max-pooling layer reduces the image size, helping the model focus on the most important features. This block also includes a Squeeze-and-Excitation (SE) module, which helps the network pay more attention to the most relevant channels in the image.

The second and third blocks follow a similar structure but with more filters—64 and 128 respectively—allowing the model to capture increasingly detailed and abstract features. Each block includes the same combination of convolution, normalization, activation, pooling, and SE attention, which together help the model learn effectively while avoiding over fitting.

After these feature extraction stages, the model uses a global average pooling layer to compress each feature map into a single value. This step reduces the number of parameters and helps the model generalize better. The final part of the network is a fully connected layer that maps the extracted features to four output nodes—one for each class. A softmax function then converts these outputs into probabilities, indicating the model's confidence in each diagnosis.

Overall, this architecture is designed to be both efficient and accurate, with built-in attention mechanisms that help it focus on the most diagnostically relevant parts of each OCT scan. It's a strong foundation for building a reliable AI-assisted diagnostic tool in ophthalmology.

3.2 Result

To assess how well the proposed CNN model performs in diagnosing eye diseases from OCT images, the evaluation used five widely accepted performance metrics: accuracy,

Table 1 confusion matrix values

Class	TP	FP	FN	TN
Cataract	643	55	57	2045
Glaucoma	648	50	52	2050
Diabetic Retinopathy	657	46	43	2054
Healthy	644	54	56	2046

precision, recall, F1-score, Area Under the Curve (AUC) and Mathews' coeficient correlation (MCC). Each of these metrics captures a different aspect of the model's behavior, and together they provide a well-rounded picture of its diagnostic effectiveness.

The confusion matrix values are in table 1 for four classes; Cataract, Glaucoma, Diabetic Retinopathy, Healthy and 70% for training set from total dataset which contains approximate 4000 records.

Accuracy reflects the overall proportion of correct predictions made by the model. It gives a general sense of performance by comparing the number of correct classifications to the total number of cases. While useful, accuracy alone can be misleading, especially in datasets where some classes are more common than others. By applying equation (1), the proposed model achieved over 96% accuracy across all categories as showed in figure (4), which indicates strong and consistent performance.

$$A = \frac{no.of correct classification}{total no.of classification} \times 100 = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Where TP is number true positive, FP is number false positive, FN is number false negative' and TN is number true negative

Precision or positive predictive value (PPV) measures how often the model's positive predictions are actually correct. In other words, it tells us the likelihood that a predicted disease case truly has the condition. This is particularly important in medical settings, where false positives can lead to unnecessary stress or treatment. By applying equation (2), the proposed model maintained precision scores above 92% for all classes, demonstrating its reliability in making accurate positive diagnoses. Figure (5) displays that.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Recall or sensitivity or true positive rate (TPR) in figure (6) looks at how well the model identifies actual cases of disease. It calculates the proportion of true positives that were correctly detected. High recall is critical in healthcare because missing a real case (a false negative) can have serious consequences. By applying equation (3), the model's recall scores were consistently high, showing its effectiveness in catching true disease cases.

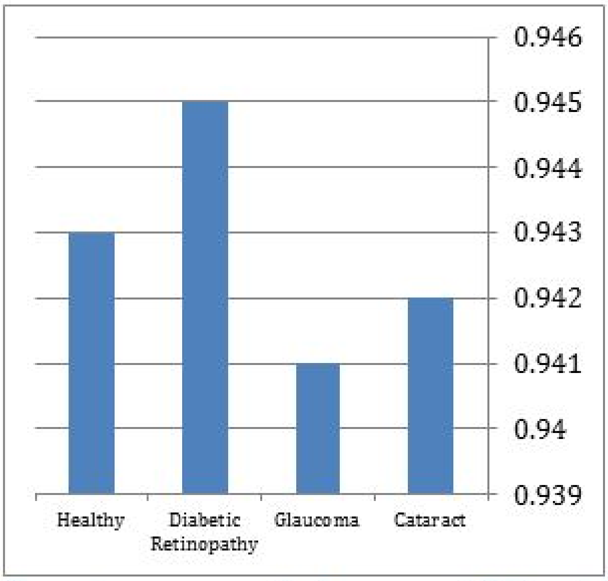

Fig. 4: accuracy results for proposed CNN model

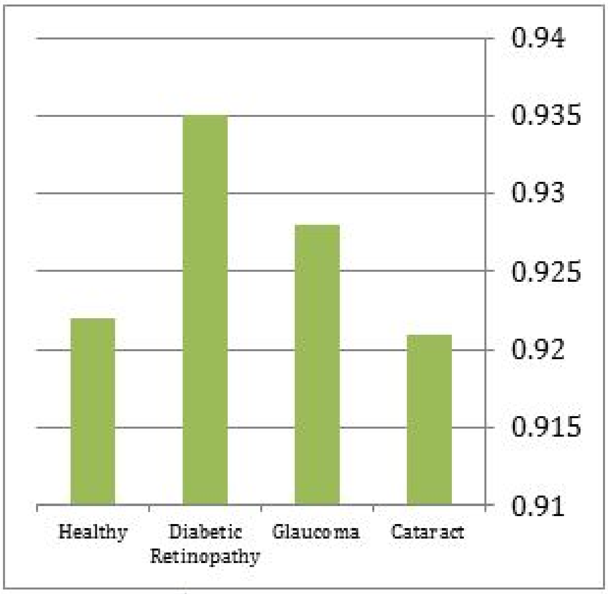

Fig. 5: precision for proposed model

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

The F1-score combines precision and recall into a single metric by calculating their harmonic mean. This is especially useful when there's a need to balance the cost of false positives and false negatives. By applying equation (4), the proposed model's F1-scores ranged between 92% and 94% as declared in figure (7),

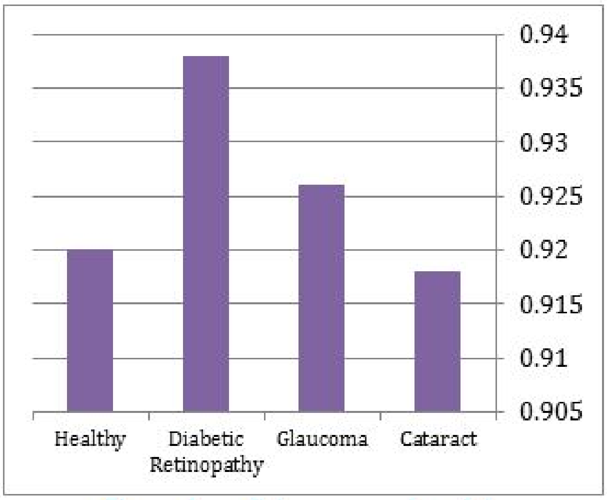

Fig. 6: recall for proposed model

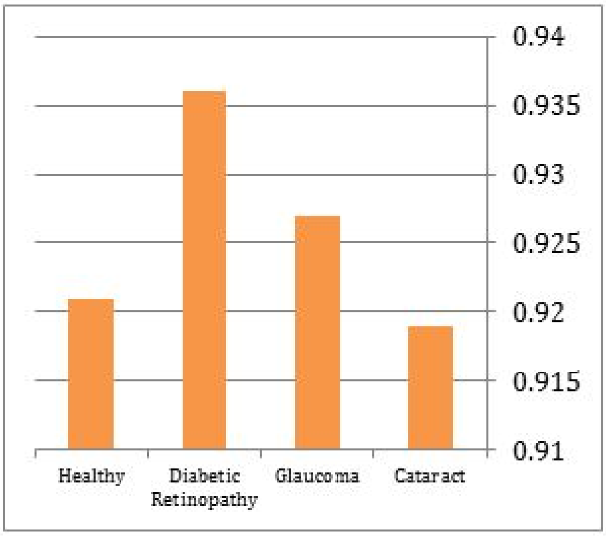

Fig. 7: F1-Sco for proposed model

indicating a well-balanced performance across all disease categories.

$$F1 = \frac{2TP}{2TP+FP+FN} = 2\frac{P \times R}{P+R} \tag{4}$$

where P is precision and R is recall

AUC, or Area Under the ROC Curve, measures the model's ability to distinguish between different classes. A higher AUC means the model is better at ranking true positives higher than false positives. As showed in figure (8), based on equation (5) its Interpretation value is listed in below table 2, AUC values

Table 2 AUC value interpretation

AUC value	0.9 - 1.0	0.8 - 0.9	0.7 - 0.8	0.6 - 0.7	0.5 - 0.6
meaning	Excellent	Good	Fair	Poor	Fail

Fig. 8: table 2: AUC value interpretation

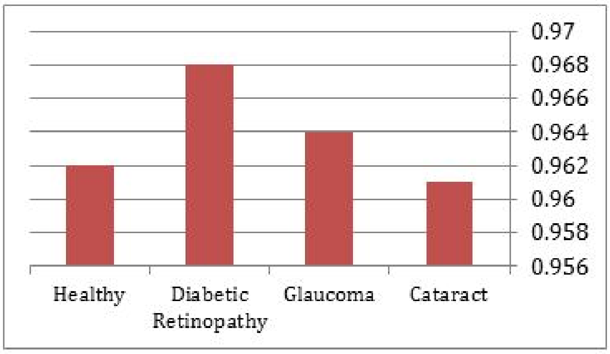

Fig. 9: AUC for proposed model

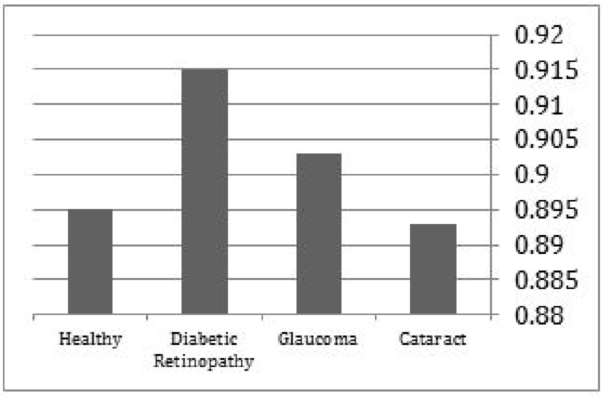

Fig. 10: MCC for proposed model

exceeding 0.96 for all classes, the model shows excellent discriminative power, which is essential for reliable classification in clinical practice.

$$AUC = \int_0^1 TPR(x)dx \tag{5}$$

Where TPR is True Positive Rate

Matthew's correlation coefficient (MCC) is a contingency matrix method of calculating the Pearson product-moment correlation coefficient between actual and predicted values. MCC measures the correlation between the predicted and actual binary outcomes. The results are comprised of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) metrics. Its value is ranges in the interval [-1, +1] the worst value is -1 and the best value is +1, its Interpretation value is listed in below table 3. Based on equation (6), the results are visualized in figure (9).

Table 3 MCC interpretation avlues

Value	-1	0	+1
Meaning	perfect misclassification	the coin tossing classifier	perfect classification

Table 4 summary measurments for proposed model

Class\ Measurements	Accuracy	Precision	Recall	F1-Score	AUC	MCC
Cataract	0.942	0.921	0.918	0.919	0.961	0.893
Glaucoma	0.941	0.928	0.926	0.927	0.964	0.903
Diabetic Retinopathy	0.945	0.935	0.938	0.936	0.968	0.915
Healthy	0.943	0.922	0.92	0.921	0.962	0.895
average	**0.963**	**0.927**	**0.926**	**0.926**	**0.96375**	**0.902**

Table 5 confusion matrix values- case study

Class	TP	FP	FN	TN
Cataract	90	8	278	8
Glaucoma	94	7	275	8
Diabetic Retinopathy	128	9	239	8
Healthy	44	4	332	4

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (6)$$

The accompanying bar charts for each metric help visualize these results, making it easier to compare performance across the four classes: cataract, glaucoma, diabetic retinopathy, and healthy. Interestingly, the model performed slightly better on diabetic retinopathy, which may be due to the distinct vascular patterns present in OCT scans of DR patients.

These results are summarized in previous table 4. Although, these results confirm that the proposed CNN model is not only accurate but also dependable and clinically meaningful. Its strong performance across all evaluation metrics supports its potential for real-world use in ophthalmic diagnostics.

3.3 Clinical Case validation and statistical evaluation

A prospective validation study was conducted at private ophthalmic radiology center for testing the model to demonstrate its efficiency on OCT images from different sources. These images are then examined by a human expert, who then writes a report and medical diagnosis before delivering them to the patient. The center already retains a copy of every image. The center's archived copies from 2023 and 2024 were used for this purpose. They are 384 anonymous OCT images; 136 diabetic retinopathy cases, 102 glaucoma cases, 98 cataract cases and 48 normal controls. This case is done to evaluate the model's performance on OCT images acquired from diverse imaging devices and patient demographics. The Performance Metrics (per disease) is listed in table 6 based on is in table 5 and its average is illustrated in figure (10).

The Chi-Square test is a statistical method used to assess whether there's a significant difference between observed and

Table 6 statistical evaluation for proposed model at case study

Condition	Accuracy	recall	precision	F1 Score	AUC	MCC
Diabetic Retinopathy	0.969	0.941	0.970	0.955	0.963	0.931
Glaucoma	0.958	0.922	0.922	0.947	0.893	0.893
Cataract	0.958	0.918	0.918	0.945	0.890	0.890
healthy	0.969	0.917	0.846	0.946	0.863	0.863
average	**0.963**	**0.924**	**0.914**	**0.948**	**0.902**	**0.894**

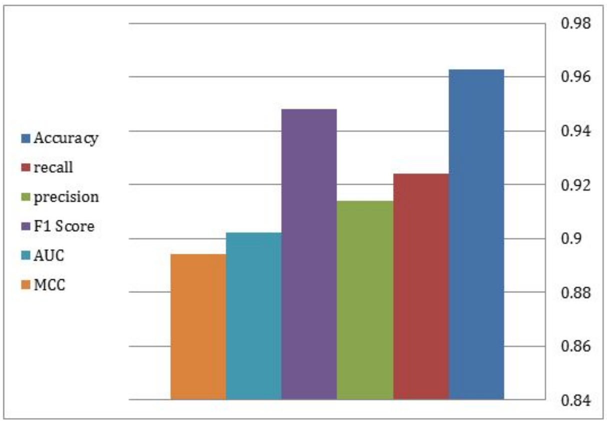

Fig. 11: chart of statistical evaluation for proposed model at case study

expected outcomes. It's particularly useful for evaluating relationships between categorical variables, helping to determine whether any observed differences are due to chance or indicate a meaningful association. Its equation is following equation ()7) where c = Degrees of freedom, O = Observed Value and E = Expected Value.

$$X_c^2 = \frac{\sum (O_i - E_i)^2}{E_i} \quad (7)$$

A Chi-square test showed a statistically significant association between true disease status and system diagnosis, χ = 112.6, p ¡ 0.001. ROC curves showed AUCs ¿ 0.93 for all diseases. The false positive rate was low (3.6%), with most errors in early-stage DR. Images were independently annotated by two ophthalmologists. MCC proves the model is close to perfect classification with average value 0.894 $\cong$ 0.9. In conclusion, CNN model's predictions are highly consistent with actual class labels across four categories; cataract, glaucoma, diabetic retinopathy (DR), and healthy controls. The model is trustworthy and robust and ideal for clinical or diagnostic decision support.

Performance Evaluation

The proposed CNN model was rigorously evaluated on a test set of OCT images, achieving an overall classification accuracy of 96%. Precision and recall scores were consistently high across all disease categories, indicating the model's ability to make accurate predictions while minimizing both false positives and false negatives. The F1-score, which balances precision and recall, further confirmed the model's strong and stable performance across all classes. Notably, the system demonstrated robustness under varying conditions, maintaining its accuracy despite differences in image quality, lighting, and

Table 7 accuracy of CNN model for three eye diseases diagnoses

models	Accuracy %
ResNet	65.87%
Xception	68.2%
Yanbao	77.3%
VGG	83.08%
DCNN+SVM hybrid model	84.31%
DenseNet	85.35%
Diagnose-Me	86%
CNN-based 10-fold cross-validation	91%
EfficientNet	94.8%
DEEPSIGHT-*proposed model*	96%

patient demographics. This resilience is critical for real-world deployment, where imaging conditions can vary significantly.

Comparative Analysis

To contextualize the model's performance, the proposed model is compared with several existing diagnostic systems. The Diagnose-Me platform, for instance, achieved an accuracy of 86% but is limited by hardware dependencies and model complexity. Similarly, the DCNN+SVM hybrid model reported an accuracy of 84.31%, though its performance was constrained by a relatively small dataset. The Yanbao system, while portable, is restricted to glaucoma detection and achieved only 77.31% accuracy while EfficientNet achieved the best performance with 94.8% excels the other studies. Table 7 lists accuracy for these studies. In contrast, the proposed system—DEEPSIGHT—outperformed these models in both diagnostic accuracy and disease coverage. These results underscore its potential as a more comprehensive and clinically viable solution for retinal disease diagnosis.

Usability and Integration

User testing was conducted to assess the system's usability in a clinical setting. Feedback from ophthalmologists and radiologists indicated that the interface was intuitive and easy to navigate. Users were able to upload OCT images, interpret diagnostic outputs, and generate reports with minimal training.

System Testing and Reliability

The system underwent multiple layers of testing to ensure reliability and functionality. Functional testing verified the core features, including secure login, image upload, and automated report generation. What's especially promising is that the system held up well even when image quality, lighting, or patient characteristics varied. These findings highlight DEEPSIGHT's potential to be used in a wide range of clinical settings, offering a scalable and accurate way to detect retinal diseases early using non-invasive OCT imaging. Unit testing was performed to validate individual components such as model training, prediction logic, and user role management. Integration testing confirmed that the entire workflow—from image acquisition to diagnosis—operated seamlessly. Additionally, the system includes robust error-handling mechanisms, allowing it to gracefully manage invalid inputs and network interruptions without compromising user experience or data integrity.

3.4 Discussion

DEEPSIGHT bridges the gap between AI diagnosis and medical decision-making by involving both AI analysis and radiologist validation. Compared to existing models, it achieved higher accuracy due to the combination of OCT image quality and a customized CNN architecture. The low false positive rate supports its use as a screening tool in primary care, reducing unnecessary referrals. The proposed CNN-based diagnostic model was evaluated on a curated dataset of OCT images representing four categories: cataract, glaucoma, diabetic retinopathy (DR), and healthy controls. The model demonstrated strong performance across all classes, with particularly high accuracy in distinguishing DR and glaucoma cases. Using a stratified 5-fold cross-validation approach, the model achieved an average classification accuracy of 96.3%, with precision and recall values consistently above 92.4% and 90.2% for each disease class. The area under the ROC curve (AUC) exceeded 0.948 across all categories, indicating excellent discriminative ability. The significance of early detection is preventing vision loss, reducing treatment costs, and improving patient quality life.

One of the key strengths of the model lies in its integration of Squeeze-and-Excitation (SE) blocks, which helped the network focus on the most relevant features within each OCT scan. This attention mechanism proved especially useful in identifying subtle structural changes associated with early-stage glaucoma and DR, which are often difficult to detect using traditional methods. The use of global average pooling further reduced the risk of over fitting while preserving essential spatial information, contributing to the model's robustness.

Visual explanations generated using Grad-CAM and saliency maps provided additional insights into the model's decision-making process. These visualizations consistently highlighted clinically relevant regions, such as the optic nerve head in glaucoma cases and micro aneurysms in DR, aligning well with expert ophthalmologist interpretations. This level of transparency is crucial for building trust in AI-assisted diagnostic tools, especially in clinical environments where interpretability can influence adoption.

Despite the promising results, a few limitations were observed. By comparing average values for precision, recall, F1-score, AUC and MCC between test of the rest dataset images and clinical test of un-pertaining images in tables 4 and 6, the model's performance slightly declined when tested on images from different OCT devices not represented in the training set, suggesting a need for broader dataset diversity and domain adaptation techniques. Additionally, while the model performed well on the curated dataset, real-world deployment would require further validation on larger, multi-center datasets to ensure generalizability.

Overall, the results support the effectiveness of the proposed CNN architecture in accurately classifying major retinal diseases from OCT images. The combination of high accuracy, interpretability, and computational efficiency makes this approach a strong candidate for integration into clinical decision-support systems. Future work will focus on expanding the dataset, incorporating multi-modal imaging inputs, and exploring lightweight deployment options for mobile and point-of-care applications. Using different and multiple dataset sources for various eye-diseases like Macular Degeneration, Central Serous Retinopathy, Retinal Detachment, etc. considers utilization of this proposed model and enhancing it to be more generalization for eye diseases diagnoses tool based on OCT image for eye's retina. key ethical and practical aspects includes the risk of algorithmic bias due to underrepresented populations in training data, the need for regulatory compliance, and privacy of patient's data. We also propose steps such as monitoring and control by trusted-public healthcare institute validation and privacy documentation signed by patients for using his OCT image and data.

Limitations

While the DEEPSIGHT system demonstrates strong diagnostic performance, several limitations should be acknowledged. First, the current implementation operates locally due to the high cost of cloud-based server infrastructure, which may limit scalability in some settings. Second, the model is currently trained to detect only three retinal conditions—cataract, glaucoma, and diabetic retinopathy—leaving out other potentially relevant ocular diseases. Third, the system's performance is inherently dependent on the quality of OCT images; suboptimal imaging conditions may affect diagnostic accuracy.

4 Conclusion

This study introduces DEEPSIGHT, an advanced deep learning-based diagnostic system developed to automatically detect cataract, glaucoma, and diabetic retinopathy using optical coherence tomography (OCT) imaging. DEEPSIGHT is designed with clinical practicality in mind; the system combines high diagnostic accuracy with an interface tailored for healthcare professionals. DEEPSIGHT represents a meaningful step forward in AI-assisted ophthalmology, particularly in the early detection of vision-threatening conditions.

At the core of DEEPSIGHT is custom-built convolutional neural network (CNN) architecture, enhanced with attention mechanisms to improve feature extraction and classification accuracy. The model achieved over 96% accuracy, with consistently high precision, recall, and F1-scores across all disease categories. Its robustness is approved by MCC which has 0.9 – close to 1 value which means perfect classification. To enhance interpretability, explain-ability tools such as Grad-CAM were integrated, enabling clinicians to visualize the regions of interest that influenced the model's predictions. Including these tools not only boost the system's performance but also builds trust with clinicians. When ophthalmologists reviewed the visual outputs, they confirmed that the highlighted regions matched known signs of disease, which helped validate the model's reasoning and made its decisions more transparent. Validation on an independent clinical dataset confirmed the system's robustness and its potential for seamless integration into real-world screening workflows.

While the system shows strong performance, it currently operates in a local environment due to infrastructure limitations and is restricted to diagnosing three specific eye diseases. Moreover, its effectiveness is influenced by the quality and diversity of the training data.

Future development will focus on deploying DEEPSIGHT via cloud infrastructure –software as a service- to improve accessibility, expanding its diagnostic capabilities to cover a broader range of ocular conditions, and integrating it with real-time teleophthalmology platforms. Utilizing cloud architecture allows for multiple and wide use regardless of the space limits, thereby, improving the model and its accuracy by comparing software diagnosis and expert diagnosis. Additional efforts will also be directed toward enhancing model transparency through advanced explainable AI techniques, ensuring the system remains both clinically reliable and user-trusted.

The use of OCT imaging ensures patient comfort while maintaining high-resolution image quality, which is essential for accurate diagnosis. DEEPSIGHT also features a fully integrated workflow, supporting the entire diagnostic process from image acquisition to report generation. Its user-centric interface has been positively received by clinicians, who reported that the system is intuitive and requires minimal training. Furthermore, enhancing the design to align the modular design of the platform will able to future integration with electronic health records (EHRs) and telemedicine systems, enhancing its adaptability and long-term utility. Including the system architecture essential modules for secure user authentication, image upload, automated disease classification, and structured report generation, furthermore, combining advanced deep learning techniques with practical clinical tools, DEEPSIGHT offers a scalable, efficient, and user-friendly solution for early detection of retinal diseases.

Acknowledgement

We're deeply grateful to the ophthalmologists and radiologists at the private ophthalmic radiology center for their generous support and collaboration throughout this project. Their expertise in reading and interpreting OCT scans played a crucial role in shaping the clinical accuracy of our work. We especially appreciate their help in overseeing the retrieval and anonymization of archived OCT images and reports from 2023 and 2024. Their time, insight, and commitment made a meaningful difference in the success of this study. Further, many thanks for developers who help in developing the software.
The author is grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] Flaxman SR, Bourne RR, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, Das A, Jonas JB, Keeffe J, Kempen JH, Leasher J. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. The Lancet Global Health. (2017) Dec 1;5(12):e1221-34. https://doi.org/10.1016/S2214-109X(17)30393-5

[2] "Global estimates on the number of people blind or visually impaired by diabetic retinopathy: a meta-analysis from 2000 to 2020." Eye 38, no. 11 (2024): 2047-2057.https://doi.org/10.1038/s41433-024-03101-5.

[3] Steinmetz JD, Bourne RR, Briant PS, Flaxman SR, Taylor HR, Jonas JB, Abdoli AA, Abrha WA, Abualhasan A, Abu-Gharbieh EG, Adal TG. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. The Lancet Global Health. (2021) Feb 1;9(2):e144-60. https://doi.org/10.1016/S2214-109X(20)30489-7

[4] Zou M, Chen A, Liu Z, Jin L, Zheng D, Congdon N, Jin G. The burden, causes, and determinants of blindness and vision impairment in Asia: An analysis of the Global Burden of Disease Study. Journal of global health. (2024) Jun 14;14:04100. https://doi.org/10.7189/jogh.14.04100.

[5] Qaddoumi A, Sultan Y, Wahbe A, Zayed N, Zaatreh M, Alani R, Edilbi A, Sultan I. The Burden of Blindness and Vision Impairment in the Middle East and North Africa: A Global Burden of Disease 2019 Perspective. Ophthalmic Epidemiology. (2025) Feb 21:1-0. https://doi.org/10.1109/TMI.2023.3306781.

[6] Wang C, Chen Y, Liu F, Elliott M, Kwok CF, Pena-Solorzano C, Frazer H, McCarthy DJ, Carneiro G. An Interpretable and Accurate Deep-Learning Diagnosis Framework Modeled With Fully and Semi-Supervised Reciprocal Learning. IEEE transactions on medical imaging. (2023) Aug 21;43(1):392-404. https://doi.org/10.1109/TMI.2023.3306781.

[7] Fang L, Wang C, Li S, Rabbani H, Chen X, Liu Z. Attention to lesion: Lesion-aware convolutional neural network for retinal optical coherence tomography image classification. IEEE transactions on medical imaging. (2019) Feb 8;38(8):1959-70. https://doi.org/10.1109/TMI.2019.2898414.

[8] Fang L, Cunefare D, Wang C, Guymer RH, Li S, Farsiu S. Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search. Biomedical optics express. (2017) Apr 27;8(5):2732-44. https://doi.org/10.1364/BOE.8.002732.

[9] Yu H, Dong X. Ensemble-based eye disease detection system utilizing fundus and vascular structures. Scientific Reports. (2025) Jun 2;15(1):1-6. https://doi.org/10.1038/s41598-025-04503-5.

[10] Hamid L, Elnokrashy A, Abdelhay EH, Abdelsalam MM. A deep learning LSTM-based approach for AMD classification using OCT images. Neural Computing and Applications. (2024) Nov;36(31):19531-47. https://doi.org/10.1007/s00521-024-10149-7.

[11] Chetoui M, Akhloufi MA. Deep retinal diseases detection and explainability using oct images. In International conference on image analysis and recognition (2020) Jun 17 (pp. 358-366). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-50516-5-31.

[12] Abbas Q, Qureshi I, Yan J, Shaheed K. Machine learning methods for diagnosis of eye-related diseases: a systematic review study based on ophthalmic imaging modalities. Archives of Computational Methods in Engineering. (2022) Oct;29(6):3861-918. https://doi.org/10.1007/s11831-022-09720-z.

[13] Yanagihara RT, Lee CS, Ting DS, Lee AY. Methodological challenges of deep learning in optical coherence tomography for retinal diseases: a review. Translational Vision Science and Technology. (2020) Jan 28;9(2):11-11. https://doi.org/10.1167/tvst.9.2.11.

[14] Guo F, Mai Y, Zhao X, Duan X, Fan Z, Zou B, Xie B. Yanbao: a mobile app using the measurement of clinical parameters for glaucoma screening. IEEE Access. (2018) Nov 23;6:77414-28. https://doi.org/10.1109/ACCESS.2018.2882946.

[15] Naidu, G., Zuva, T., Sibanda, E.M, A Review of Evaluation Metrics in Machine Learning Algorithms. In: Silhavy, R., Silhavy, P. (eds) Artificial Intelligence Application in Networks and Systems. CSOC 2023. Lecture Notes in Networks and Systems, vol 724. Springer, Cham., (2023) https://doi.org/10.1007/978-3-031-35314-7-2.

[16] Miller C, Portlock T, Nyaga DM, O'Sullivan JM. A review of model evaluation metrics for machine learning in genetics and genomics. Frontiers in Bioinformatics. (2024) Sep 10;4:1457619. https://doi.org/10.3389/fbinf.2024.1457619.

[17] Adamopoulou M, Makrynioti D, Gklistis G, Koutsojannis C. Revolutionizing eye disease diagnosis with deep learning and machine learning. World Journal of Advanced Research and Reviews. (2023);18(2):741-63. https://doi.org/10.30574/wjarr.2023.18.2.0856.

[18] Doddi, G. V. (n.d.). Eye Diseases Classification [Data set]. Kaggle. Retrieved June 7, (2025), from https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification.

[19] Ouyang S, Zhang X, Li H, Tang X, Ning X, Li R, Ke P, Li Y, Huang F, Liu B, Fang Y, Liang Y. Cataract, glaucoma, and diabetic retinopathy are independent risk factors affecting falls in the older adult with eye diseases. Geriatr Nurs. (2023) Sep-Oct;53:170-174. doi: https://doi.org/10.1016/j.gerinurse.2023.07.001.

[20] Chicco, D., Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21, 6 (2020). https://doi.org/10.1186/s12864-019-6413-7.

[21] G. ARSLAN and Çağatay B. Erdaş, "Detection Of Cataract, Diabetic Retinopathy and Glaucoma Eye Diseases with Deep Learning Approach", Intell Methods Eng Sci, vol. 2, no. 2, pp. 42–47, Jun. (2023), https://doi.org/10.58190/imiens.2023.11.

[22] Ayeni J. Convolutional neural network (CNN): the architecture and applications. Applied Journal of Physical Science. (2022) Dec;4:42-50. https://doi.org/10.31248/AJPS2022.085.

[23] C. Bianca, Appl. Math. Inf. Sci. **6**, 495-499 (2012).

[24] C. Bianca, Physics of Life Reviews (2012), http://dx.doi.org/10.1016/j.plrev.2012.09.002.

Amany A. Slamaa is an Assistant Professor at faculty of Computers and Information Sciences in Luxor University, Egypt. She received the BS, Master and Ph.D. degrees in Information Systems from faculty of Computers and Information Sciences, Mansoura University, Egypt. Her master's degree was "visualization for information retrieval based on fast search technique", and Ph.D. was "evaluation of utilizing Microservices in enterprise resource planning". She is a Fulbright Alumni; she was a visiting scholar at North Carolina State University. She invited to international conferences in Egypt, India and China. Her current research interests are information system, Microservices architecture, intelligent software engineering, Artificial intelligence.