FUDMA Journal of Engineering and Technology
ISSN: 3092-9385
https://fjet.fudutsinma.edu.ng/index.php/fjet/
Volume 2, Issue 2, 2026, 45-52

Development of an Ocular Disease Prediction Using Deep Learning

Toyin OKEBULE

Department of Computer Science, Faculty of Computing, Federal University of Technology and Environmental Sciences, Iyin-Ekiti, Nigeria

slimtoyin2003@gmail.com

Abstract

Ocular diseases remain a leading cause of visual disability, significantly affecting quality of life and productivity, particularly in low- and middle-income countries with limited access to specialized eye care. Early and accurate diagnosis relies heavily on experienced ophthalmologists, despite advancements in imaging techniques such as fundus photography, optical coherence tomography (OCT), and ultrasound. This study presents a hybrid deep learning framework for multi-class ocular disease classification, leveraging Convolutional Neural Networks (CNNs) for spatial feature extraction and ensemble methods for classification refinement. The dataset employed consists of 7,230 ocular images across five disease categories. The dataset includes patients aged 30–80 years, with 55% male and 45% female, predominantly of Nigerian African ethnicity. All images are fundus photographs (no OCT or multimodal images), ensuring consistent imaging modality. The dataset consists of diabetic retinopathy (1,750 images), glaucoma (1,420), macular degeneration (1,260), cataracts (1,100), and normal images (1,700). Extensive preprocessing including resizing, normalization, contrast enhancement and data augmentation was applied to enhance model robustness and generalization. The model was trained and validated using stratified splits (70% training, 15% validation, 15% testing). The ensemble approach outperformed individual models, achieving 97.8% accuracy, with a classification report confirming minimized false positives and false negatives, critical in clinical diagnostics. The model also demonstrated low inference time and high computational efficiency, supporting potential deployment in clinical decision-support systems and mobile diagnostic applications.

Keywords: *Ocular disease, Convolutional Neural Network, DenseNet, ResNet, Data Preprocessing.*

1.0 Introduction

Ocular diseases such as diabetic retinopathy, glaucoma, age-related macular degeneration (AMD), cataracts, and ocular tumors are among the leading causes of vision impairment and blindness globally, affecting over 250 million adults [1]. These conditions often progress silently, with patients remaining asymptomatic until significant damage has occurred. Enhancing accessibility to screening and diagnostics is thus critical for improving global eye health. This underscores the necessity for early detection, yet many regions, particularly in low- and middle-income countries, face severe shortages of ophthalmologists and lack adequate screening infrastructure [2]. Early identification and treatment are crucial for reducing the effects of these diseases; however, obstacles like inadequate access to eye care professionals, high medical expenses, and uneven healthcare infrastructure, particularly in under-resourced regions continue to impede prompt detection [3]. Traditionally, the diagnoses of ocular diseases rely on clinical evaluation and interpretation of retinal images such as fundus photographs, optical coherence tomography (OCT) scans, and fluoresce in angiography. These procedures demand significant time, expert input, and consistent accuracy, which may not always be feasible in real-world healthcare environments. The growing availability of ocular image datasets and the advancement of artificial intelligence (AI) technologies, particularly deep learning, offer a transformative opportunity to bridge this diagnostic gap [4]. Furthermore, the lack of interpretability in deep learning systems is one of the major concerns addressed in this study. In clinical settings, physicians are reluctant to trust “black-box” AI models that do not offer visual or logical reasoning for their outputs. Techniques like Grad-CAM and SHAP have been proposed to enhance transparency, but they are not yet widely integrated into multi-disease ocular diagnostic frameworks [5]. Fairness in AI predictions remains an unresolved issue, as many models fail to account for demographic variations such as age, gender, and ethnicity, resulting in unequal diagnostic outcomes. In recent years, deep learning has revolutionized medical imaging, offering remarkable improvements in the detection and classification of ocular diseases from sources such as fundus photographs, optical coherence tomography (OCT), ultrasound, and slit-lamp imaging. Convolutional Neural Networks (CNNs) in particular have shown exceptional performance [6]. Despite the promise of CNN's approach, ensemble learning approaches that combine different deep learning architectures with traditional models are gaining traction for ocular disease classification. Hybrid techniques like stacking, which merge CNN feature extraction with classifiers such as XGBoost, have boosted model accuracy to more than 90%, thereby improving clinician trust through interpretable visual explanations and achieving high diagnostic performance [7]. However, several challenges remain in this field, including domain shifts due to varied imaging hardware and the lack of population diversity in most datasets. Techniques of domain adaptation, such as Collaborative Feature

Received: 27/05/2026; Revised: 06/06/2026; Accepted: 09/06/2026; Published: 13/07/2026

Ensembling Adaptation (CFEA), have helped mitigate these issues, enhancing cross-device reliability [8]. This study aims to develop an ocular disease prediction model using deep learning by providing an accurate, automated, and scalable model that enhances early detection, reduces diagnostic errors, and supports ophthalmologists with faster decision-making. The model's integration of fairness and interpretability ensures equitable and trustworthy healthcare delivery across diverse populations.

In order to gain insight into this study, several related works and methods were thoroughly researched, including:

Deepak et al. [9] proposed an optimized convolutional neural network (CNN) leveraging deep learning techniques for the multi-class classification of eye diseases, including normal vision, glaucoma, and cataracts. The research involves fine-tuning three pre-trained CNN models (SqueezeNet, Darknet-53, and EfficientNet-b0) by adjusting optimizer algorithms (SGDM, RMSProp, and Adam) and varying batch sizes (6, 8, and 10). Among the models, Darknet-53 achieved the highest accuracy of 99.4% on a test set of 1,000 images when configured with a batch size of 6 and the Adam optimizer. A confusion matrix was employed to evaluate the model's performance metrics. In a comparative analysis, the highest accuracies recorded were 95% for SqueezeNet, 99.4% for Darknet-53, and 90% for EfficientNet-b0. Nevertheless, the study is constrained by the use of a limited and uniform dataset, which could reduce the applicability of its findings to broader, real-world clinical environments. Another study in [10] experimented with a real-time eye disease detection system that uses ensemble deep learning to categorize diabetic retinopathy, cataracts, glaucoma, and pathological myopia. To increase accuracy, the model combines ensemble methods with CNNs. 14,400 fundus photos were collected from various sources, including the Ocular Disease Recognition dataset, JSIEC, and Kaggle-based datasets on diabetic retinopathy and glaucoma. The best performance was achieved by the stacking model with 95% accuracy, while voting achieved 94%, boosting achieved 93%, and bagging achieved 87% accuracy. The strength in their models lies in the capacity to manage image variability and enhance diagnosis accuracy for a variety of illnesses. However, the study's performance in practical applications may be impacted by inconsistent image quality and a lack of clinical validation. One study in [11] utilized a deep learning system for classifying ocular diseases with various features, including diabetic retinopathy, cataracts, glaucoma, and normal eyes, using a 4,200-image dataset of retinal fundus images sourced from Kaggle and the Indian Diabetic Retinopathy Image Dataset (IDRiD). The study compared two deep learning models: the Convolutional Neural Network (CNN) and the Neural Network (NN) model based on EfficientNet. The NN model outperformed the CNN with 94% accuracy compared to the CNN's 84% accuracy. The EfficientNet model also recorded a high F1 score of 99%, particularly in detecting diabetic retinopathy. Their study highlights the effectiveness of image-based deep learning models in improving automated eye disease detection, particularly in settings with limited access to specialists. Future work includes incorporating more disease classes and optimizing the system for mobile deployment. Another paper in [12] proposed eye diseases prediction and classification using deep learning approach with Gabor filter as a feature extractor from the images. The images contained 4217 with four different classes which include Diabetic Retinopathy, Cataract, Glaucoma and Normal were implemented on Convolution Neural Network models (VGG-16). The experimental result shows that VGG-16 achieved accuracy of 85.34%.

Abdullah et al. [13] presented a comprehensive review of deep learning (DL) techniques for the automated detection and classification of common ocular diseases, including glaucoma, cataracts, diabetic retinopathy (DR), myopia, and age-related macular degeneration (AMD). The study emphasized the importance of early detection using CNN-based models such as VGG16, ResNet, and Inception, which have shown high accuracy in identifying multiple eye diseases simultaneously. The study reviewed publicly available fundus image datasets and discussed the general pipeline of disease classification, including preprocessing and the use of pre-trained CNNs for feature extraction. Moreover, it highlighted the limitations, such as the lack of theoretical grounding in existing DL models, challenges in hardware demands, and generalizability across populations. However, they proposed future research directions involving hybrid models that combine CNNs with ensemble learning methods to improve classification performance, reduce computational cost, and support real-time deployment in clinical settings.

Gulati et al. [14] proposed a deep learning-based ensemble model for the multiclass classification of ocular disorders using Optical Coherence Tomography (OCT) images. The ensemble approach integrated features extracted from pre-trained models—VGG16, Xception, and MobileNet were fed into a CNN classifier. The model was trained and evaluated on a publicly available Kaggle OCT retinal dataset comprising 84,495 images across four classes: Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), Drusen, and Normal. Among the ensemble configurations, the MobileNet+CNN model achieved the highest performance, with an average classification accuracy of 95.34%. However, the study did not implement broader classification, nor did it explore model generalization using larger and more diverse datasets from real-world clinical environments.

Another study in [15] implemented the Ocular Disease Detection System (ODDS) using a deep learning-based diagnostic tool to detect multiple ocular disorders in retinal fundus images (RFI). The EfficientNet-B3 architecture is used by the system to classify images. Training, validation, and testing were conducted using a carefully selected dataset of 5,600 fundus photos from five different disease categories. The study's findings: ODDS achieved a 93%

Received: 27/05/2026; Revised: 06/06/2026; Accepted: 09/06/2026; Published: 13/07/2026

testing accuracy rate with proper preprocessing. The approach has great promise for use in rural healthcare settings with limited ophthalmological resources and in community screening. However, there is a need to improve the model's generalization capability and the dataset's size. Sara Ejaz et *al.* [16] proposed a deep learning-based diagnostic framework for the detection of multiple ocular diseases using color retinal fundus images. The study utilized the publicly available Retinal Fundus Multi-Disease Image Dataset (RFMiD). Preprocessing techniques were applied to enhance model performance. Experiments were conducted using two CNN architectures, notably a 12-layer and a 20-layer CNN. The CNN model achieved a validation accuracy of 90.34% and a testing accuracy of 89.59%, though it exhibited signs of overfitting. The study demonstrated that deep CNNs could effectively learn discriminative features from fundus images for multi-disease classification.

One study in [17] utilized a deep learning-based approach for multi-disease classification of retinal fundus images using a novel Multilevel Glowworm Swarm Convolutional Neural Network (MGSCNN) by classifying 39 different retinal conditions using the Kaggle sourced dataset, the Retinal Fundus Multi-Disease Image Dataset (RFMiD), which contains approximately 3,200 images. The model achieved impressive performance metrics, including an accuracy of 95.09%, sensitivity of 96.59%, specificity of 93.59%, precision of 93.53%, recall of 96.67%, and F1-score of 95.07%. The study demonstrated the potential of the MGSCNN model in improving the accuracy and efficiency of automated multi-disease detection in ophthalmology. However, this study did not implement an ensemble learning approach.

2.0 Materials and Methods

A. Research and Summary

This section elaborates on the procedures used for training and validating the deep learning models, including the choice of architecture, the rationale behind using certain hyper parameters, and the performance metrics employed in model building, as shown in Figure 1:

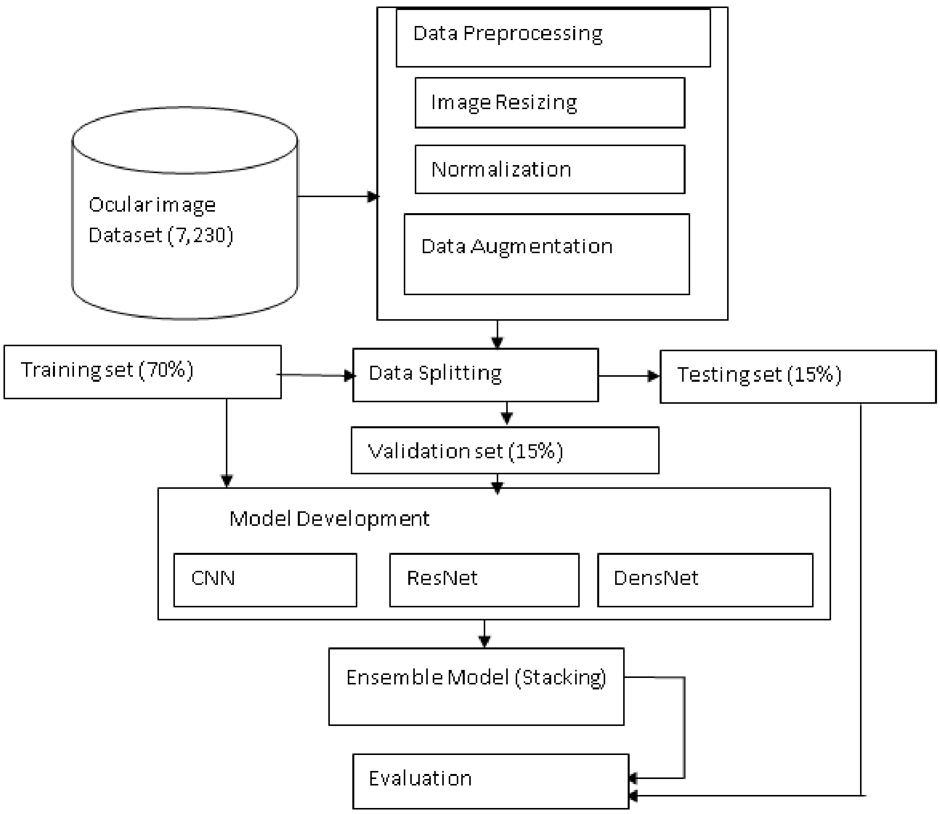

Figure 1: The System's Development Block Diagram

B. **Data Preprocessing**

The dataset shows moderate class imbalance, which reflects real-world ophthalmic screening distributions which includes patients aged 30–80 years. This demographic representation ensures the dataset's relevance to local ophthalmic populations. To mitigate this bias, data augmentation techniques were applied, including rotation of ±15°, horizontal and vertical flipping, zooming (set between 0.9–1.1), and brightness adjustments (set between 0.8–1.2), particularly for underrepresented classes. Moreover, class-weighted training was used during model

Received: 27/05/2026; Revised: 06/06/2026; Accepted: 09/06/2026; Published: 13/07/2026

optimization to promote balanced learning across all disease categories. Table 1, shows the Class-wise Distribution of the Ocular Disease Dataset.

Table 1: Class-wise distribution of the ocular disease dataset

Ocular Condition	Number of Images	Percentage (%)
Diabetic Retinopathy	1,750	24.2
Glaucoma	1,420	19.6
Macular Degeneration	1,260	17.4
Cataract	1,100	15.2
Normal (Healthy)	1,700	23.6
Total	7,230	100.0

The following Data Preprocessing was carried out on Ocular dataset employed for this study:

a. **Image resizing:** To ensure uniformity in model input and enhance training efficiency, all images were resized to 224×224 pixels.

b. **Data augmentation:** To enhance model robustness and generalization, various data augmentation techniques were applied to the training images. Each image underwent random horizontal and vertical flipping with a probability of 0.5, rotation within a range of ±15°, and shearing transformations up to 10°. Additionally, zooming was applied within a scale range of 0.9–1.1, and brightness adjustments were applied with a factor ranging from 0.8–1.2.

c. **Normalization:** Normalization was applied to all input images. Pixel values, originally in the range [0, 255], were scaled to [0, 1] by dividing each pixel by 255. This helps the model converge faster during training and reduces the risk of numerical instability.

d. **Data splitting:** To evaluate model generalization, the dataset was split into 70% for training, 15% for validation, and 15% for testing. This approach enabled effective learning, parameter tuning, and unbiased assessment of model performance on unseen images, further strengthening generalization and reducing the risk of overfitting.

C. Model Development

The models utilized for this study are explained as follows:

a. Convolutional Neural Networks (CNNs) Model

One of the deep learning models employed in this study is the Convolutional Neural Network (CNN), which is highly effective for ocular disease prediction using image-based data [18]. The CNN model was implemented on 7,230 retinal fundus images. Each image is processed through a series of convolutional layers, where filters automatically learn to extract key visual features present in the dataset. These layers are followed by ReLU activation functions, which introduce non-linearity and enhance the model's ability to capture complex patterns. Pooling layers, particularly max pooling, are then used to reduce the dimensionality of the features, preserving the most significant information while decreasing computational complexity and minimizing the risk of overfitting. The resulting output is flattened and passed through fully connected (dense) layers that learn deeper feature relationships to make final predictions. The last layer uses a Softmax activation function to generate a probability distribution across various ocular disease classes, indicating the most likely diagnosis for each input image. The model is trained using back propagation on labeled data, and with proper preprocessing and parameter tuning, the CNN outperformed the other independent deep learning models used in the study, achieving an accuracy rate of 97%. This makes it a robust and suitable model for ocular disease prediction using a Softmax-based output layer,

$$\hat{y} = Softmat(W_0 . \phi(\mathrm{Flatten}(\mathrm{Pool}(\sigma(\mathrm{Conv}(\mathrm{x}))))) + b_o \quad (1)$$

where (x) is the retinal fundus image, (Conv) is the Convolutional operation that applies filters to extract features, σ represent the activation function Introduced to ReLU, b_o represents the bias term of the output layer, Pool represent the operation that down samples max pooling, $\mathrm{Flatten}$ function transforms the feature maps into a vector, (ϕ) is the dense layer(s) for complex representations, $Output\ layer$ Computes probability distribution over classes using Softmax, $\hat{y}$ represent the Output prediction vector with class probabilities

b. ResNet Model

In this study, the ResNet architecture was applied on 7,230 retinal fundus images for ocular disease prediction. The ResNet model has the advantage of utilizing residual connections, which allow very deep networks to be trained without performance degradation. Unlike the AlexNet or VGG model that uses large filters and is prone

Received: 27/05/2026; Revised: 06/06/2026; Accepted: 09/06/2026; Published: 13/07/2026

to overfitting, ResNet employs stacked 3×3 convolutional layers with skip connections to preserve feature information across layers. The model includes ReLU activations, max pooling (2×2), and fully connected layers, concluding with a Softmax layer for ocular disease classification. The ResNet model achieved better accuracy on the RFMiD dataset, demonstrating strong generalization and robust performance in detecting multiple ocular diseases. The ResNet model function is explained in Equation 2:

$$\hat{y}\_ResNet = Softmax(W_0 . \phi(\mathrm{Flatten}(\mathrm{F(x)})) + b_0) \quad (2)$$

where F(X) is the residual block defined as:

$$\mathrm{F(x)} = \sigma(\mathrm{x} + \mathrm{F}(\mathrm{X}, \{W_0\}))) \quad (3)$$

where $\mathrm{F}(\mathrm{x}, \{W_0\})$ represent the transformation via stacked 3×3 convolutional layers, $\mathrm{x} + \mathrm{F}(.) represents$ skip connection, which is the residual learning, σ is the activation function ReLU and b_o represents the bias term of the output layer.

c. **DenseNet Model**

DenseNet was applied to 7,230 retinal fundus images of ocular disease for multi-class ocular disease classification. In this study, DenseNet was connected to all the previous layers, thereby allowing and improving feature reuse and gradient flow. Each dense block contains repeated 1×1 and 3×3 convolutions, while transition layers with 1×1 convolutions and 2×2 average pooling reduce feature map size to prevent overfitting. The DenseNet-121 architecture achieved better performance by capturing fine-grained retinal features and ensuring balanced classification across disease classes. The mathematical function of the DenseNet model is represented as:

$$\mathrm{x}\ell = \mathrm{H}\ell([\mathrm{x}^0, \mathrm{x}^1, \ldots, \mathrm{x}\ell - 1]) \quad (4)$$

where $\mathrm{x}\ell$ is the output of the $\mathrm{x}\ell - \mathrm{th}$ layer, $\mathrm{H}\ell$ is a composite function of operations (ReLU), $[\mathrm{x}_0, \mathrm{x}_1, \ldots, \mathrm{x}\ell - 1]$ represents the concatenation of feature maps from all previous layers

d. **Ensemble Learning** Techniques

In this study, a stacked ensemble learning approach was employed to improve the accuracy and robustness of ocular disease prediction using retinal fundus images. The ensemble approach combined the three powerful Convolutional neural network architectures: CNN, ResNet, and DenseNet with each trained on a dataset of 7,230 images. Each model produced class probabilities for the input images, and these outputs were then passed to a meta-classifier in a stacking ensemble framework. The ensemble approach was employed in this study to reduced independent model biases, enhanced generalization across multiple ocular disease classes, and demonstrated superior performance compared to independent model approaches. Ensemble Classifier Prediction is explain as follows:

Let x be the input retinal fundus image, $M_1, M_2, and\ M_3$, represent the three base models**:** CNN**,** ResNet**,** and DenseNet**,** respectively,
$p_1(x), p_2(x)\ and p_3(x)$, represent the class probability vectors (from Softmax outputs) predicted by each model for input dataset. H represents the meta-classifier that learns to combine these predictions (stacking ensemble) $\hat{y}$represent the final predicted class label. Therefore, the Base Model Predictions:

$$z = [(p_1(x), p_2(x), dp_3(x))] \quad (5)$$

where each $[p_i(x) \in \mathbb{R}^C$ with C being the number of classes
$z \in \mathbb{R}^{3C}$ represent the concatenated vector of all base model outputs.
Meta-Classifier Prediction:

$$\hat{y} = H(z) = (p_1(x), p_2(x) p_3(x)) \quad (6)$$

where H is the stacking ensemble, that optimally combine the predictions from CNN, ResNet, and DenseNet.

e. **Model Interpretability and Clinical Explainability**

Grad-CAM and SHAP techniques were used to visualize the retinal regions driving ensemble model predictions. Grad-CAM highlighted disease-specific features such as micro-aneurysms for diabetic retinopathy, optic nerve cupping for glaucoma, drusen for macular degeneration, and lens opacities for cataracts. SHAP provided quantitative feature importance, confirming that the model focused on clinically relevant pathological areas. This interpretability aligns predictions with ophthalmological knowledge, enhancing clinician trust and supporting diagnostic adoption.

f. **Evaluation Metrics**

The key metrics employed in this study are defined as follows:
Accuracy measures the proportion of total correct predictions

$$\boldsymbol{Accuracy} = \frac{\mathbf{TP+TN}}{\mathbf{TP+TN+FP+FN}} \tag{7}$$

Precision assesses the correctness of positive predictions:

$$\boldsymbol{Precison} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+FP}} \tag{8}$$

Recall (Sensitivity) quantifies the model's ability to detect true positives:

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \tag{9}$$

F1-Score, the harmonic mean of precision and recall:

$$\boldsymbol{F1} = \mathbf{2}.\frac{\boldsymbol{Precision\ .Recall}}{\boldsymbol{Precision+Recall}} \tag{10}$$

AUC-ROC evaluates the trade-off between the true positive rate (TPR) and false positive rate (FPR) across different thresholds.

To validate the generalizability and robustness of the model, k-fold cross-validation was employed. This method divides the dataset into *k* subsets (folds), where the model is iteratively trained on *k - 1* folds and tested on the remaining fold. The average of the results across all folds provides a more reliable estimate of real-world performance and reduces variance caused by random train-test splits.

$$Sensivity = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negative\ (FN)} \tag{11}$$

$$False\ Positive\ Rate\ (FPR) = \frac{False\ Positive\ (FP)}{False\ Positive\ (FP) + True\ Negative\ (TN)} \tag{12}$$

The AUC is computed by integrating the ROC curve:

$$AUC = \int_0^1 TPR(FPR) fFPR \tag{13}$$

Typically calculated using numerical approximation methods like trapezoidal rule.

$$Average\ Accuracy - \frac{1}{k}\sum_{j-1}^{k} Accuracyi \tag{14}$$

where TP = True Positives, TN = True Negatives, FP = False Positives and FN = False Negatives

3.0 Results and Discussion

Implementation was carried out using Python programming Language in Google Colab environment. The Machine Learning models were implemented and their performances were evaluated using Scikit-learn - a Python Library. The evaluation of the three Deep Learning models' performances is displayed through confusion matrices, from which values were extracted to calculate accuracy, precision, recall, AUC-ROC and F1-scores. The result is presented in the Table 2.

Received: 27/05/2026; Revised: 06/06/2026; Accepted: 09/06/2026; Published: 13/07/2026

Table 2: Performance comparison of the model with 95% confidence intervals

Metric	CNN	ResNet	DenseNet	Ensemble
Accuracy (%)	95.2	94.5	92.6	97.8
95% CI (Accuracy)	94.3 – 96.1	93.5 – 95.5	91.5 – 93.7	97.2 – 98.4
AUC-ROC	0.930	0.940	0.914	0.990
95% CI (AUC-ROC)	0.920 – 0.940	0.930 – 0.950	0.902 – 0.926	0.985 – 0.995
McNemar's Test (p-value, vs. Ensemble)	$p < 0.05$	$p < 0.05$	$p < 0.05$	Not Applicable
Inference Time	Low	Moderate	High	Low

Quantitatively, Table 2 presents the comparative performance, statistical significance, and inference time of the models evaluated. The CNN model achieved an accuracy of 95.2% (95% CI: 94.3–96.1) with an AUC-ROC of 0.930 (95% CI: 0.920–0.940), demonstrating strong baseline performance and low inference time. ResNet followed closely with an accuracy of 94.5% (95% CI: 93.5–95.5) and an AUC-ROC of 0.940 (95% CI: 0.930–0.950), exhibiting moderate inference time and competitive discriminative capability. DenseNet recorded an accuracy of 92.6% (95% CI: 91.5–93.7) and an AUC-ROC of 0.914 (95% CI: 0.902–0.926). However, its dense connectivity resulted in higher inference time despite maintaining stable classification performance. The ensemble model, which integrates predictions from CNN, ResNet, and DenseNet, delivered the best overall results with an accuracy of 97.8% (95% CI: 97.2–98.4) and an AUC-ROC of 0.990 (95% CI: 0.985–0.995), while maintaining low inference time. McNemar's statistical test confirmed that the performance improvements of the ensemble model over all individual models were statistically significant ($p < 0.05$). Overall, the ensemble consistently outperformed the standalone architectures across key evaluation metrics, demonstrating enhanced robustness, generalization, and suitability for real-world ocular disease screening and clinical decision-support applications.

4.0 Conclusion

In this study, a comprehensive deep learning approach was implemented for ocular disease prediction with 7,230 annotated fundus images, incorporating it with preprocessing, model training, and ensemble learning. Preprocessing steps included image resizing to 224×224 pixels, normalization to [0, 1], and extensive data augmentation techniques such as flipping, rotation, and brightness adjustments to enhance model generalization. The dataset was split into 70% training, 15% validation, and 15% testing. The study employed three deep learning convolutional neural network architectures—CNN, ResNet, and DenseNet. These were trained independently and as an ensemble. The CNN model achieved a strong performance with 95.2% accuracy. The ResNet model achieved 94.5% accuracy, while the DenseNet model achieved 92.6% accuracy. A stacked ensemble model was then created by combining the Softmax outputs of all three models into a new model that learned to best combine their predictions. This ensemble outperformed the independent models with a predictive performance of 97.8% accuracy, demonstrating improved robustness and generalization over individual models. Evaluation metrics such as precision, recall, F1-score, and AUC-ROC further validated model effectiveness. Additionally, k-fold cross-validation was adopted to ensure reliability and minimize variance. Our study achieved an improved accuracy of 97.8% using an ensemble model approach, demonstrating better robustness and generalization compared to previous studies in [12], which reported 87% accuracy using a lightweight CNN, and [16], which achieved 90.34% validation accuracy with CNN architectures. This shows the potential of this study's approach for accurate and reliable diagnosis of various diseases. Furthermore, the ensemble approach effectively leveraged the strengths of each architecture, enhancing classification accuracy and making it a promising solution for real-world ocular disease diagnosis. In order to assess the model's effectiveness in actual clinical contexts, future research can concentrate on augmenting the dataset, examining other ensemble approaches, integrating with clinical decision support systems, and conducting prediction studies.

Acknowledgements

The author is grateful to the management of Federal University of Technology and Environmental Sciences, Iyin-Ekiti, Nigeria for providing a research environment for this work.

References

[1] P. Vashist, S. S. Senjam, V. Gupta, N. Gupta, B. R. Shamanna, M. Wadhwani, P. Shukla, S. Manna, S. Yadav, and A. Bharadwaj, "Blindness and visual impairment and their causes in India: Results of a nationally representative survey," PLoS One, vol. 17(7), p. e0271736, 2022: https://doi.org/10.1371/journal.pone.0271736.

[2] P. Y. A. Attafuah, P. Mordi, E. K. Aggrey, A. S. P. Quaicoe, and J. O. Sarfo, "Prevalence and management of cataracts among older adults in Sub-Saharan Africa: a scoping review," BMC Ophthalmology, vol. 24, pp. 434, 2024, doi: 10.1186/s12886-024-03701-5.

[3] A. Hoagland and S. Kipping, "Challenges in Promoting Health Equity and Reducing Disparities in Access Across New and Established Technologies," Canadian Journal of Cardiology, vol. 40, no. 6, pp. 1154–1167, 2024, doi: 10.1016/j.cjca.2024.02.014.

[4] H. Bilal, A. Keles, and M. Bendechache, "Advances in disease detection through retinal imaging: A systematic review," Computers in Biology and Medicine, vol. 194(110412), 2025, doi: 10.1016/j.compbiomed.2025.110412.

[5] P. Onebunne and B. Alade, "Bias and fairness in AI models: Addressing disparities in machine learning applications," International Research Journal of Modernization in Engineering Technology and Science, vol. 6(9), pp. 1921–1934, Sep. 2024,: https://doi.org/10.56726/IRJMETS61692.

[6] M. Adamopoulou, D. Makrynioti, G. Gklistis, and C. Koutsojannis, "Revolutionizing eye disease diagnosis with deep learning and machine learning," World Journal of Advanced Research and Reviews, vol. 18(1), pp. 741–763, 2023, doi: 10.30574/wjarr.2023.18.2.0856.

[7] J. Aslam, M. A. Arshed, S. Iqbal, and H. M. Hasnain, "Deep learning based multi-class eye disease classification: Enhancing vision health diagnosis," Scientific Reports*, vol. 14(1), pp. 1–14, 2024. doi: 10.1038/s41598-024-72237-x.

[8] A. Musa, R. Prasad, and M. Hernandez, "Addressing cross-population domain shift in chest X-ray classification through supervised adversarial domain adaptation," *Scientific Reports*, vol. 15, pp. 11383, 2025. doi: 10.1038/s41598-025-95390-3.

[9] D. Deepak and S. K. Bhat, "Deep learning-based CNN for multiclassification of ocular diseases using transfer learning," *J. Med. Imaging Health Inform.*, vol. 12, no. 10, pp. 1–3, 2024, doi: 10.1080/21681163.2024.2335959.

[10] K. Sruthi, R. Sandhiya, L. M. Jeevika, A. Kirsa, S. Lakshmipraba, and T. Sukirtha, "Real-time detection of ocular diseases using ensemble deep learning," *Preprint*, SSRN, 2024.: https://ssrn.com/abstract=5089005.

[11] T. Babaqi, M. Jaradat, A. E. Yildirim, S. H. Al-Nimer, and D. Won, "Eye disease classification using deep learning techniques," in Proc. IISE Annu. Conf. & Expo, 2023.: https://doi.org/10.48550/arXiv.2307.10501

[12] A. Fahmy, "Eye Diseases Prediction and Classification using Deep Learning Techniques," Annals of Biostatistics & Biometric Applications, vol. 6(2), pp. 1–14, 2024. doi: 10.33552/ABBA.2024.06.000634.

[13] A. Aldhahab, and H. M. Al Abboodi, "Review of eye diseases detection and classification using deep learning techniques," *International Journal of Health Sciences*, vol. 6, no. S1, pp. 13064–13078, 2022.

[14] S. Gulati, K. Guleria, and N. Goyal, "Detection and Multiclass Classification of Ocular Diseases using Deep Learning-based Ensemble Model," *International Journal of Intelligent Systems and Applications in Engineering (IJISAE)*, vol. 12, no. 19s, pp. 18–29, 2024.

[15] P. Thiagarajan and M. Suguna, "Deep Learning Ocular Disease Detection System (ODDS)," in Proc. 9th Int. Conf. Mining Intelligence and Knowledge Exploration (MIKE 2023), Kristiansand, Norway, pp. 213–224, 2023. doi: 10.1007/978-3-031-44084-7_21.

[16] S. Ejaz, R. Baig, Z. Ashraf, M. M. Alnfiai, M. M. Alnahari, and R. M. Alotaibi, "A deep learning framework for the early detection of multi-retinal diseases," PLOS ONE, vol. 19 (7), e0307317, Jul. 2024. https://doi.org/10.1371/journal.pone.0307317.

[17] R. Chavan and D. Pete, "Automatic multi-disease classification on retinal images using multilevel glowworm swarm convolutional neural network," *Journal of Engineering and Applied Science*, vol. 71 (26) 2024: https://doi.org/10.1186/s44147-023-00335-0.

[18] T. Okebule, "CNN-Based Classification for Lumpy Skin Disease Detection (Bovine-HealthGuard)," Iconic Research and Engineering Journals, vol. 8(12), pp. 966–974, 2025.