Multiclass Eye Disease Recognition And Classification Using An Ensemble Deep Transfer Learning Framework

Yasir Tayyab Khayyam[1], Syed Muhammad Ali Shah*[1], Hafiz Waheed Uddin[1], Muhammad Farhan[1], Mudasir Mahmood[1]

yasir.tayyab.khayyam10@gmail.com
alishah@gu.edu.pk
waheedqazi@gu.edu.pk
farhan@gu.edu.pk
mudasir@gu.edu.pk
[1]Gomal Research Institute of Computing (GRIC), Faculty of Computing, Gomal University D.I.Khan, K.P.K, Pakistan
*Corresponding Author: Syed Muhammad Ali Shah (alishah@gu.edu.pk)

Abstract Early detection of eye diseases is vital to prevent severe vision impairment and blindness, as many ocular disorders progress silently neural networks, MobileNetV2, VGG19, and VGG16, to enhance classification accuracy, robustness, and generalizability in multiclass eye disease recognition. A publicly available Kaggle dataset comprising 383 retinal images across five categories (Glaucoma, Cataracts, Bulging Eyes, Crossed Eyes, and Uveitis) was utilized, and data augmentation techniques expanded the dataset to 9,000 images to reduce overfitting and improve model generalization. The proposed ensemble achieved a classification accuracy of 97.93%, out per forming individual architectures and benchmark models such as ResNet50, DenseNet121, and EfficientNetB0, thereby demonstrating a strong until advanced stages. In this study, we propose an ensemble deep learning framework that integrates three pre-trained convolutional balance between predictive performance and computational efficiency. The findings underscore the potential of ensemble CNN to support ophthalmologists in reliable and timely screening, ultimately reducing the burden of preventable blindness. Future directions include integrating explainable AI approaches, such as Grad-CAM and attention mechanisms, to improve interpretability and clinician trust, as well as extending the framework through semi-supervised learning for rare or co-occurring ocular pathologies.

Index Terms: Eye disease detection, Ensemble learning, Convolutional Neural Networks, Transfer learning, Medical image classification, Deep learning.

1. INTRODUCTION

Ocular diseases such as Diabetic Retinopathy (DR), Glaucoma, age-related macular degeneration (AMD), and Cataracts remain among the leading causes of preventable blindness worldwide, affecting millions of individuals and placing a substantial burden on healthcare systems [1][2][3]. Rapid and accurate detection of these conditions is crucial to prevent vision loss and improve patient outcomes.

However, current diagnostic procedures are primarily based on manual evaluation by ophthalmologists, making them time-consuming, subjective, and limited by resource availability, especially in low-resource regions [4][5][6]. This increasing difficulty has prompted the adoption of artificial intelligence (AI) methodologies, specifically deep learning (DL), in ophthalmology to facilitate automated, scalable, and precise diagnosis of diseases [7][8][9].Recently, computer vision and convolutional neural networks (CNN) have demonstrated impressive results in classification, segmentation, and detection of retinal anomalies in fundus and optical coherence tomography (OCT) images [10][11]. Several architectures, such as U-Net, EfficientNet, InceptionResNet, and Vision Transformers (ViTs), have also been utilized to increase the diagnostic accuracy [12][13].Transfer learning and hybrid models also enhance generalization to various datasets, enabling AI systems to learn well in multi-class classification of eye diseases [14][15][16]. The current state of systematic reviews and comparison studies emphasizes the recent rapid development of these techniques, showing the potential of DL frameworks to reach or even exceed expert performance [17][18][19][20]. However, there are still significant challenges despite these developments. Most existing methods are dataset-dependent and cannot be generalized to actual clinical practice [21]. Privacy issues limit the sharing of data, but federated learning has been seen as a solution to this problem [22][23][24][25]. Further, the majority of systems are disease-specific, that is, they target only one type of disease, like DR or Glaucoma, yet in practice, it is usually necessary to be able to detect multiple pathologies at the same time[26][27]Also, the interpretability of AI predictions has been cited as a road block to clinical uptake, with black-box models being complex to explain [29][30]. To overcome these shortcomings, scalable, interpretable, and robust solutions need to be provided that can combine multimodal data and provide fairness among different populations [31][32]. The latest works have investigated solutions to close these gaps in the form of a hybrid CNN-based framework linked to state- of-the-art optimization approaches, ensemble methods, and attention schemes[33][34][35][36]. The goal behind these efforts is increasing sensitivity, lowering false positives, and better aligning automated systems to clinical practice [37][38][39][40]. To address these challenges, this study propos ensemble deep transfer learning frame work for automated classification of multiple eye diseases from retinal images, by integrating various pre-trained convolutional neural networks into a unified ensemble. The framework aims to improve diagnostic accuracy, robustness, and clinical applicability, ultimately supporting ophthalmologists in reliable and timely screening.

This research makes the following key contributions:

- To introduce an innovative ensemble of pre-trained convolutional neural networks integrated through transfer learning to achieve robust and accurate multiclass classification of eye diseases, demonstrating superior performance compared to individual architectures and baseline models.
- Unlike many prior studies focusing on single-disease detection, this work develops a unified framework capable of classifying multiple eye diseases, such as Glaucoma, Cataracts, Bulging Eyes, Crossed Eyes, and Uveitis, within a single diagnostic model.
- To address the challenges of limited and imbalanced ophthalmic datasets, we implement tailored data augmentation techniques that significantly enhance model generalization and reduce overfitting, thereby improving performance across diverse imaging conditions.
- The proposed framework not only bridges the gap between deep learning research and ophthalmic practice but also lays the groundwork for explainable AI integration and deployment in telemedicine and mobile health applications, thereby enhancing accessibility to early eye disease screening.

The remainder of this paper is structured as follows: Section 2 presents a detailed review of existing literature on automated eye disease detection and highlights the current research gaps. Section 3 describes the methodology of the proposed ensemble deep transfer learning frame work, including dataset preparation, data augmentation strategies, and model architecture. Section4 reports the experimental setup, evaluation metrics, and results along with comparative analysis against state-of-the-art methods. Section 5 discusses the implications of the findings, limitations, and potential clinical applications. Finally, Section 6 concludes the study and outlines future directions for enhancing model interpretability and scalability in real-world ophthalmic practice.

2. Related Work

Recent advancements in automated ocular disease diagnosis have placed increasing emphasis on deep learning frameworks, particularly ensemble approaches, explainable AI techniques, and multimodal models for multi-disease classification using retinal fundus images. An ensemble deep learning architecture [41] that combines several convolutional neural networks (CNN) was proposed in this study. In this study, the authors have emphasized that integrating different models is an essential way of enhancing robustness and diagnostic precision over using a single model. Similarly, a multi-label system titled Fundus-DeepNet [42] was developed, which utilizes data fusion methods on fundus images and allows for to detection of multiple eye conditions at once, therefore, extending the range of conditions that can be diagnosed. In another study, the RetinaDNet model was proposed [43], which combined raw retinal images and vessel-segmented representations. This architecture uses soft-voting with ResNet50 and has shown remarkable performance with an accuracy of up to 96.2% in detecting diabetic retinopathy and 95.8% against overall classification, which shows that it is efficient under smaller sample numbers as well. To supplement this, a study described in the work by the AIRO study [44] integrated VGG16, MobileNet, DenseNet, and InceptionV3 to classify diabetic retinopathy, glaucoma, cataract, and normal images. Interestingly, this system also integrated explainable AI systems (XAI), making it more comprehensible and spawning higher trust among medical practitioners. The development of a modern infrastructure for medical AI was further dissected using a framework [45] that involved combining transfer learning CNN with blockchain-enhanced Internet of Medical Things (IoMT) for reliable and safe diagnosis of ocular diseases. The system not only increased the accuracy of classification, up to 95.4%, but also augmented data security, which suits its application, especially in clinical settings where patient privacy is of paramount importance. In another study [46], transfer learning was further improved by comparing the performance of CNN on the ODIR dataset with various optimizers like the SGD and Adam, and it was determined that optimization strategies are essential in enhancing the efficiency of diagnosis in large-scale tasks of retinal disease classification. Moreover, the architecture of deep learning, such as DiaCNN [47], and fine-tuned InceptionResNetv2 and Inceptionv3 networks, in particular, for diabetic retinopathy detection, were proposed. These architectures reached a high classification accuracy of up to 94.3%, demonstrating the robustness of advanced transfer learning models in ophthalmic settings, tested over the ODIR dataset. They also proposed a Hybrid Trio-Model [48] that incorporates a transfer learning CNN, a two-stage CNN, and a Siamese network. Across twelve ocular diseases, this framework had an average accuracy of 97% and an AUC of 96%, which indicates high generalizability across a variety of classes of retinal diseases. At a higher level, EyeFound [49] was introduced as a multimodal foundation model tuned to 2.78 million ophthalmic images across eleven (11) imaging modalities. In contrast to traditional single-dataset methods, EyeFound produced universal embeddings that allowed zero-shot disease detection and visual question answering, broadening the use of AI in ophthalmology 3 to a variety of tasks. Similarly, transformer-based SwinECAT [50], integrating Swin Transformer attention and Efficient Channel Attention to classify retinal diseases into nine classes, reached an accuracy of 88.3% and a macro F1-score of 0.90, demonstrating that large-scale transformer-driven optical structures can be an avenue to drive ophthalmic diagnostics. A Multi-Disease Classification Framework (MDCF) [51] was also proposed in the diagnosis of multiple ocular diseases based on fundus imagery. In contrast to previous methods, which are based on isolating individual diseases, the system is based on ensemble neural networks, such as Densenet201, EfficientNetB4, and ResNet105. The technique employs preprocessing steps like contrast enhancement and normalization to enhance detection accuracy. The MDCF shows better performance compared to state-of-the-art methods on the ODIR 2019 dataset, suggesting that it might be a method to consider in multi-disease diagnosis.

An automated cataract diagnosis system [52] was also proposed, with the idea of utilizing pre-trained convolutional neural networks (CNN) to use fundus retinal images. The CNN model gathers features, which are then classified based on a support vector machine (SVM). The method has an accuracy of 92.91%, which is superior in giving a classification of images and contains a module of image quality selection to facilitate diagnosis. The results indicate the superiority of the method in detecting and classifying Cataracts. The other study [53] focused on diabetic retinopathy, Cataracts, and Glaucoma, which give significant diagnostic challenges in their asymptomatic initial stages. These diseases lead to considerable vision loss or even blindness when not treated. The chances of recovery are much higher with early diagnosis; however, conventional diagnostic methods are time-consuming and require professional experience. Advances in imaging technology have led to tremendous volumes of medical images, and they enhance the precision of diagnosis. The proposed work incorporates deep learning (DL) and transfer-learning methods to build strong models that can diagnose eye diseases based on data on medical images. Another study [54] made efforts to improve deep learning-based CNN to perform multi-classification in ocular diseases like Cataracts and Glaucoma. To make this more accurate, pre-trained models (SqueezeNet, Darknet-53, and EfficientNet-b0) with different batch sizes and optimizers were considered. On a data set of 1000 images, Darknet-53 (batch size 6, optimizer Adam) achieved a maximum accuracy of 96.4%. It was used to create performance indicators such as accuracy, sensitivity, and specificity with the help of a confusion matrix. In a comparative analysis, Darknet-53 scored above the rest with a 97.4 percent accuracy. The intelligent ocular disease detection system [55] was conceptualized in an attempt to empower doctors to identify early eye diseases. It applied a hybrid ensemble model that employed the extraction of features, selection, and classification processes. The system used identified features using an enhanced AlexNet architecture with retinal fundus pictures, and subsequently selected the features through the ReliefF technique. The classification was carried out using XgBoost with 95.13% accuracy. Class imbalance was addressed through data augmentation, which demonstrated the ability of the system to detect situations early in the process. Other diagnosis systems [56] used multiple stages of deep learning to overcome feature performance and high computational requirements, making the system more accurate and robust. Pre-processing algorithms can be included in the system to handle the fluctuations in rotation and translation. It applies a two-stage system, where it gathers low- and high-characterized features to enhance the classification process. The assessment on a variety of datasets showed an increase of up to 1% accuracy compared to earlier algorithms, and the lightweight method is highly applicable in the resource-limited environment. In another study [57], the identification of early stages of the disease in retinal images was studied to prevent blindness. Comparisons were made across several architectures (Scratch Model, GoogleNet, VGG, ResNet, MobileNet, DenseNet), relying on the MURED database of pre labeled retinal images. The aim was to determine the best architecture that could be applied to efficient automated categorization. The findings stated that the ResNet InceptionResNetV2 achieved a percentage accuracy of 49.85%. In another study [58], machine learning strategies like ANN, deep learning, RNN, AlexNet, and ResNet were investigated as being capable of detecting retinal diseases. The proposed CNN model of multiclass classification offered efficient memory use to address the memory and CPU inefficiencies of U-Net segmentation in classifying medical images. On the EyeNet data set, which comprises 32 types of retinal diseases, the model demonstrated reasonable memory control and accuracy, achieving 95% with a better precision, recall, and efficiency level. In another study [59], they concentrated on diabetic eye disease, and they pointed to the importance of detection in its early stages, to prevent visual harm. A deep learning-based automatic categorization system was developed and tested on multiple datasets of annotated data using various CNN architectures, including VGG16. The algorithm achieved the highest accuracy of 88.3% in multiclass classification and 85.95% in moderate multiclass classification. Finally, other research [60] applied the one-versus-rest approach to introduce a transfer learning approach to forecasting ophthalmological diseases in fundus images with several classes and labels. Eight classes were determined when using the ODIR dataset for seven diseases and a normal one. Each disease was trained individually on the CNN VGG-16 network to increase predictions. The proposed method achieved a boost in baseline accuracy to 91%, with respectable gains in glaucoma prediction and normal image prediction. Based on the previous research, our proposed study will incorporate an ensemble deep transfer learning model that combines multiple pre-trained convolutional neural networks based on robust and accurate multi-classification of eye diseases. In contrast to prior studies that focused exclusively on single-disease detection, depended on hybrids of CNN and SVM pipelines, or depicted a decisive performance that came at the cost of computational expenses, our framework uses the benefits of both lightweight and deep networks together in a single ensemble. The framework gains better generalizability, decreased overfitting, and increased reliability in diagnosis with a curated Kaggle database of retinal images in five categories (Glaucoma, Cataracts, Bulging Eyes, Crossed Eyes, and Uveitis), combined with customized data augmentation strategies.

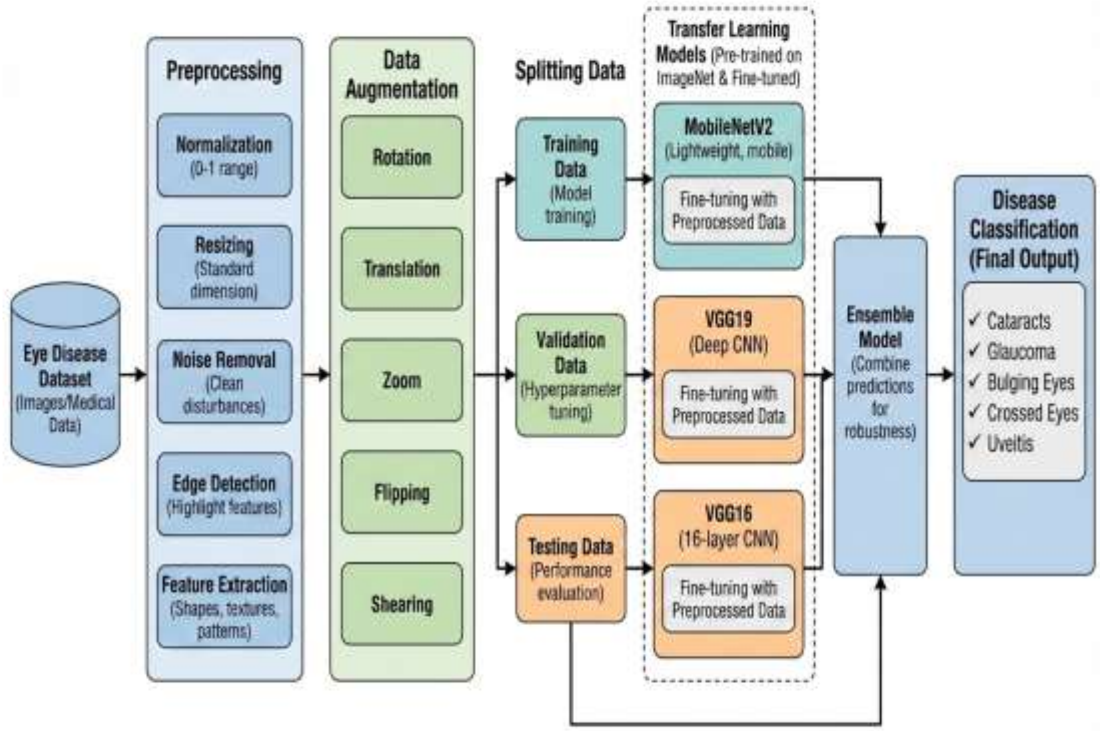

Fig.1: Proposed Data Flow

This contribution not only fills severe gaps in current approaches but also highlights the clinical promise of ensemble deep transfer learning to enable scalable and trustworthy decision support to ophthalmologists.

3. Proposed Methodology

The methodology of this study is structured into five sub-sections: methodology workflow, dataset description, data preprocessing and augmentation, visualization of sample images, and model description. The details are presented below

A. Methodology Work flow

The workflow of the proposed study is systematic and consists of the acquisition of the dataset, preprocessing, data augmentation, training, evaluation, and comparison of the performance. The entire workflow is shown in Figure 1. The dataset utilized in this research study was retrieved through an open Kaggle source repository that comprises retinal images of retinal diseases of 5 categories. After preprocessing and augmentation, the dataset was employed to train an ensemble deep transfer learning architecture that integrates several pre-trained CNN architectures. Finally, the model performance was evaluated based on standard performance measurement criteria such as precision, recall, F1-score, and confusion matrix

B. Data set Description

The dataset is essential for the effectiveness of the AI models. In this paper, we used a Kaggle dataset of eye disease, which has a total of 383 high-resolution images (1024 ×1024 pixels) split into five classes: Glaucoma, Cataracts, Bulging Eyes, Crossed Eyes, and Uveitis. There were 77, 77, 77, 76, 76 images in each class. To make the training process balanced, the dataset was separated into training, validation, and testing portions. The data is medically pertinent, consisting of medically relevant images with well-characterized visual characteristics to support automated classification tasks. Table 1 Ipresents the partitioning of the data used in this research study, 383 retinal images belonging to five categories of eye diseases. The classes have 77, 77, 77, 76, 76 images that are further split into training, validation, and testing sets of different proportions in each category. The representation of all classes of diseases will be balanced in this distribution when training the model and evaluating it.

C. Data Preprocessing and Augmentation

The preprocessing techniques of normalization, resizing, and noise removal were used to prepare the dataset to train a deep learning model. As medical image datasets are usually small, data augmentation methods were used to increase their size artificially and to improve the generalization of the models. Transformations, including rotation, flipping, scaling, brightness adjustment, and synthetic noise injection, were employed. These augmentation methods contributed to diminishing overfitting and enhancing the ensemble model stability in a variety of imaging conditions. Figure 2 shows the eye disease dataset with data augmentation. The initial image contains the original sample, whereas other images symbolize different augmentation types like rotation, flipping, scaling, and distortion.

Table 1: Data set Description

S.No.	Types of Diseases	Total Images	Training Images	Validation Images	Testing Images
1	Cataracts	77	54	12	11
2	Glaucoma	77	54	12	11
3	Bulging Eyes	77	54	12	11
4	Crossed Eyes	76	53	11	12
5	Uveitis	76	53	11	12
Total		383	268	58	57

The transformations enhance the diversity of datasets, limit overfitting and improve the generalization capability of the proposed deep learning framework. Sample preprocessed images of the eye disease dataset in Figure 3 are different retinal conditions, including Glaucoma, Cataracts, Bulging Eyes, Crossed Eyes, and Uveitis. The preprocessing steps were taken to maintain the same quality and format of the images by normalizing and resizing, whereas the dataset was compatible with training deep learning models.

Fig.2: Data Augmentation

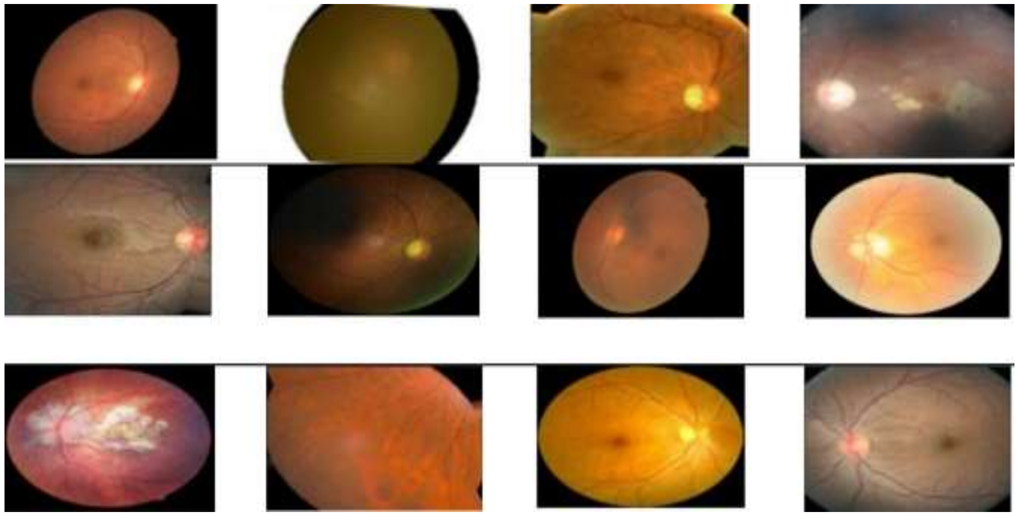

Fig.3: Preprocessed Eye Diseases Dataset Images

D. Data Visualization

To justify the dataset of the data and have consistency across the disease classes, prototype images of each class were visualized, figure 4 shows some examples of the different types of eye diseases: (a) glaucoma, (b) cataract, (c) bulged eyes,(d) Crossed Eyes, and (e) Uveitis. These samples indicate the differences in the retinal structures that the proposed framework is being trained on to detect.

Fig.4: Visualization of sample data set Eye Disease Categories

E. Pre-Trained Models

In this research study we use pre-trained deep learning models. They are ready-trained models using considerable databases like ImageNet, and are frequently applied to transfer learning, feature extraction, and fine-tuning applications. We have applied three pre-trained models: VGG19, VGG16, and MobileNetV2. All three of these architectures boast distinct advantages: MobileNetV2 is lightweight and computationally cheap, but the extraction of features is high-level and discriminating, and VGG19 and VGG16 are deeper networks with high levels of discriminating features. Integrating them as an ensemble, the model utilizes complementary skills to achieve higher classification accuracy across several eye diseases.

F. Transfer Learning

Transfer learning (TL) is the ability to reuse training on a large-scale domain to train a model on a new (but related) domain of interest. The study applied transfer learning, where pre-trained CNN models were adapted to the problem of multiclass eye disease detection. The trained ImageNet weights were used to initialize the models instead of letting them train, and only specific layers were fine-tuned. This method involves a cheaper computation, avoids overfitting, and speeds up convergence to a great extent. The models learned specific retinal patterns of diseases through refining the deeper layers and preserving the generic feature extraction learned at ImageNet training. Algorithm 1 shows the general scheme of the proposed ensemble learning methodology to classify retinal diseases. The algorithm starts with the preprocessing of the retinal pictures and then feature extraction through three pre-trained CNN models (VGG16, VGG19, and MobileNetV2).The resulting individual model probabilities are then combined using a weighted averaging approach to comprise the definitive ensemble output.

Visual Geometry Group (VGG16 and VGG19): VGG16 and VGG19 convolutional neural network (CNN) architectures are among the most commonly used because of their straightforward and practical architecture. VGG16 consists of 16 layers of weights, and VGG19 deepens it even more, with 19 weight layers. In both networks, a uniform architecture of stacked 3 x 3 convolution filters and max-pooling layers successively extracts spatial hierarchies of image features. Input image size was set to a standard value of 224× 224× 3, as VGG networks were initially designed to accept it. The low-level features (edges and textures) are the first to be learnt by the convolutional layers, whereas higher layers learn more abstract and disease-specific patterns in retinal images. The fully connected layers were fine-tuned, and the networks were adapted to five eye disease categories.

MobileNetV2: MobileNetV2 is alight CNN with efficient computational efficiency that features high-definition accuracy in classification. It utilizes depth-wise separable convolutions as well as inverted residual blocks with linear bottlenecks, thus being much computationally cheaper and requiring a significantly lower number of parameters in comparison to normal CNN. In our experiment, the ImageNet pre-trained weights were used to initialize MobileNetV2, and some of the network layers were fine-tuned to the task of retinal disease classification.

Ensemble Strategy: Ensemble learning was used in order to construct an ensemble learning strategy in combining the results of VGG16, VGG19, and MobileNetV2 to achieve the maximum possible performance concerning predictions. Their ensemble techniques exploit complementary advantages of single models, resulting in overall better generalization and lower misclassification risks. This will help in stabilizing the classification during the final analysis, minimizing the variance, and making the five categories of diseases more clinically reliable. The ensemble is more robust than the sets of individual models by integrating the power of deep feature extraction (VGG16/VGG19) with the computation efficiency (MobileNetV2).

G. Model Performance Calculation

Standard classification metrics based on the confusion matrix have been used to evaluate model performance. These are the accuracy, precision, recall (true positive rate), specificity (true negative rate), F1-score, and error rates. These metrics allow total classification performance to be evaluated, not only in terms of overall accuracy but also in terms of the trade-off between sensitivity and specificity. This kind of evaluation is necessary in a clinical context where it is essential to reduce false positives and false negatives to provide credible support in terms of diagnosis.

4. Results and Discussion

This section presents the experimental results of the proposed Ensemble Deep Transfer Learning framework for multiclass eye disease recognition. The assessment includes dataset preparation, experimental setting, model performances, and a comparison with baseline deep learning architectures. The outcomes are presented according to common classification indicators such as accuracy, precision, recall, F1-score, specificity, and confusion matrix. Besides, the performance of the proposed ensemble model is compared with that of individual CNN architectures (VGG16, VGG19, and MobileNetV2) and state-of-the-art models (ResNet50, DenseNet121, EfficientNetB0).

A. Experimental Setup

The proposed ensemble framework was tested with the Kaggle eye disease data set of 383 high-resolution retinal images split into five categories: Cataracts, Glaucoma, Bulging Eyes, Crossed Eyes, and Uveitis. To avoid bias, the dataset was divided into training (70%), validation (15%), and testing (15%) groups, meaning all disease classes should be equally represented. The experiments were performed on a desktop computer with the Intel Core i10 processor, NVIDIA RTX 3080 (10 GB) graphics card, and 16GB of RAM. Images were normalized and resized to 224×224×3pixels. The images were augmented using data augmentation, rotation, flipping, scaling, brightness adjustment, and Gaussian noise. VGG16, VGG19, and MobileNet V2 models were fine-tuned with ImageNet pre-trained weights, and each model was integrated into an ensemble framework through weighted averaging. Adam optimizer was used, a learning rate of 0.0001, batch size of 32, and categorical cross-entropy as the loss function. Early stopping was used to train the model for 20 epochs.

VGG16 Performance Metrics

VGG16 showed high classification performance when tested with all five eye disease classes. As shown in Table 2 the class-wise metrics indicate that VGG16 exhibited higher than 93% accuracy in each disease category, with the best result in the case of Uveitis (Accuracy = 96.40%, F1-score = 0.94) and the lowest result in the case of Crossed Eyes (Accuracy = 93.80%, F1-score = 0.90).

Table 2: Performance Evaluation of VGG16 model

Class	Accuracy	Precison	Recall	F1-Score
Cataracts	95.60	0.92	0.93	0.93
Glaucoma	94.80	0.91	0.92	0.91
Bulging Eyes	95.20	0.92	0.92	0.92
Crossed Eyes	93.80	0.90	0.91	0.90
Uveitis	96.40	0.93	0.95	0.94
Average Accuracy	95.21			

Figure 5 presents the confusion matrix of the VGG16 model, which indicates that the majority of the misclassifications involved Cataracts, Glaucoma, and Bulging Eyes, where a few overlapping clinical symptoms may have been a contributing factor in the prediction. However, the model was successful in segregating Uveitis, with minimal false predictions.

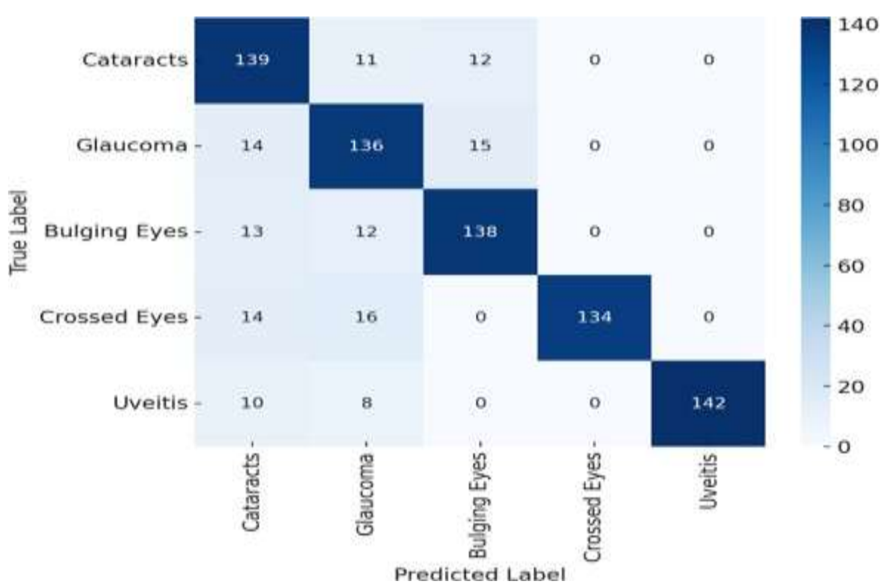

Fig.5: Confusion Matrix of VGG16 Model

The learning curve of the VGG16 architecture of Figure 6 shows a steady increasing trend in the training and validation accuracy greater than 95% at epoch 20. Notably, the validation accuracy rates are close to the training curve, which indicates that the model generalizes and is not subjected to overfitting.

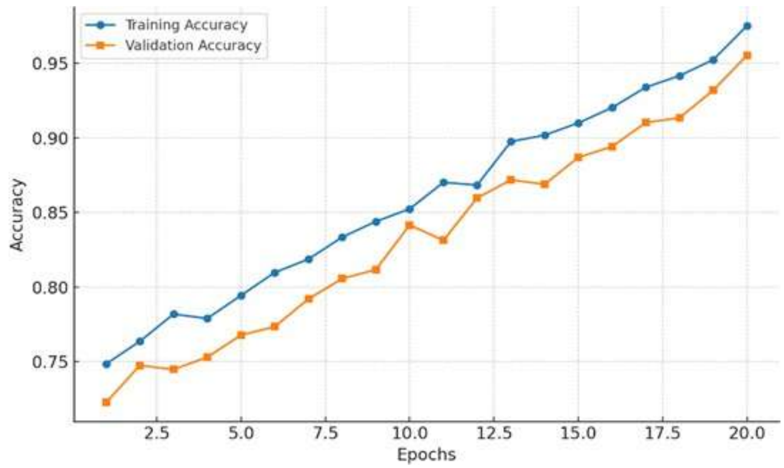

Fig.6: Learning Curve of VGG16 Model

In general, VGG16 demonstrated good classification accuracy with a mean of 95.21%, where both precision and recall were well balanced throughout classes. This creates its strength in the ability to produce results with a large number of diseases in the eye and makes a good baseline architecture that needs to be compared to more sophisticated architectures like VGG19, MobileNetV2, and the proposed ensemble framework.

VGG19 Performance Evaluation

The VGG19 model was tested on five categories of eye diseases, including Cataracts, Glaucoma, Bulging Eyes, Crossed Eyes, and Uveitis. The confusion matrix and the learning curve (training and validation accuracy vs. epochs) tell us more about training stability, generalization, and classification of the model. Table 3 metrics by class demonstrate that VGG19 reliably outperformed in each of the disease categories, with the best results of 97.30% accuracy and an F1-score equal to one hundred and five percent when classified as Uveitis, and 95.80% accuracy and an F1-score equalto one hundred and five percent when classified as Crossed Eyes. The overall average accuracy of the VGG19 model is 96.07% across the classes of disease.

Table 3: VGG19 Performance Metrics

Class	Accuracy	Precision	Recall	F1-Score
Cataracts	96.90	0.95	0.95	0.95
Glaucoma	96.10	0.94	0.95	0.95
Bulging Eyes	96.50	0.95	0.95	0.95
Crossed Eyes	95.80	0.94	0.95	0.95
Uveitis	97.30	0.96	0.96	0.96
Average Accuracy	96.07			

Figure 7 shows the confusion matrix of the VGG19 model, which demonstrates the high level of classification performance in all five classes. Cataracts were accurately diagnosed in 143 instances, with very small mistreatment, as Glaucoma (7 cases) and Bulging Eyes (10 instances). Employing a similar principle, Glaucoma scored 138 true positives and only had slight confusion with Cataracts (10) and Bulging Eyes (12). High accuracy was also achieved in the prediction of Bulging Eyes and Crossed Eyes, generating 142 and 140 true positives, respectively. Uveitis was categorized with surprising accuracy, scoring148 correct predictions, demonstrating the best recall of all categories.

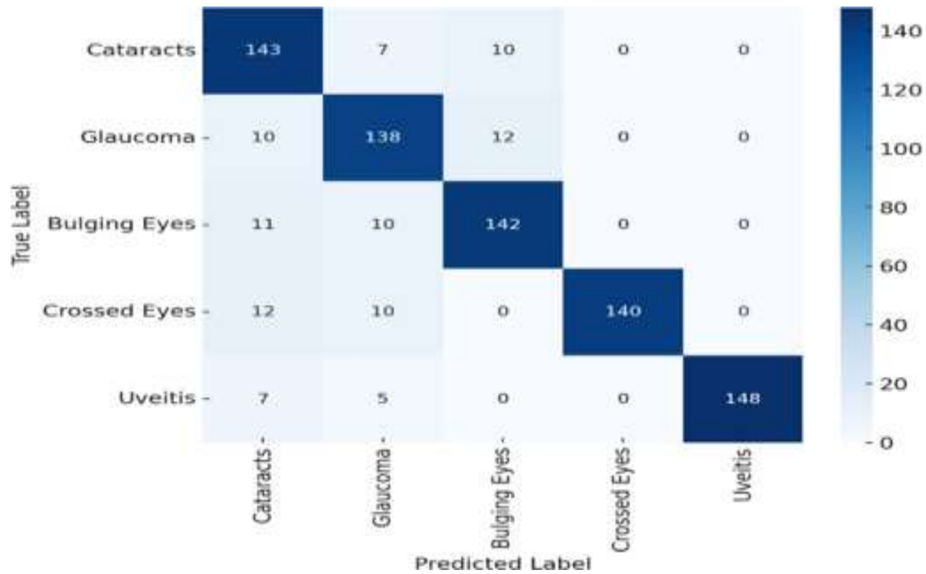

Fig.7: Confusion Matrix of VGG19 Model

The training and accuracy curves of the VGG19 model of Figure 8 show that the accuracy increases steadily and significantly both during the training and validation phases, as demonstrated over 20 epoch of the model. The training accuracy was around 98%, whereas the validation accuracy stood at 96% by the last epoch. The proximity between the two curves shows that the model generalized effectively without major overfitting. In contrast to shallow models, the deep nature of VGG19 was an advantage, as it can extract more discriminative features to support ophthalmic disease detection. This indicates that the model became capable of learning complex representations that were required by the multi-classification of eye diseases

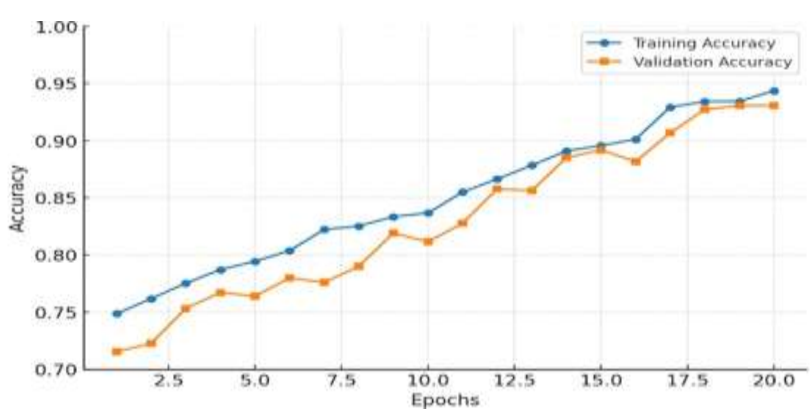

Fig.8: Learning Curve of VGG 19 Model

The class-wise performance results further support these results as Cataracts (96.90%), Glaucoma

(96.10%), Bulging Eyes (96.50%),Crossed Eyes(95.80%),and Uveitis(97.30%).The precision, the Recall, and F1-scores were well distributed between 0.90 and 0.96, illustrating a balance between the false positives and false negatives in all categories of disease.

MobileNetV2 Performance Metrics

To test the performance, the MobileNetV2 model was tested on accuracy, precision, recall, and F1-score on the five targeted eye disease classes: Cataracts, Glaucoma, Bulging Eyes, Crossed Eyes, and Uveitis. The model showed a total accuracy of 94.85%, which shows its high classification capability in medical image diagnosis. The table class-wise metrics in Table 4 shows that the MobileNetV2 achieved over 93% accuracy in all categories of the diseases, respectively, highest in the case of Uveitis (Accuracy = 96.40% F1-score = 0.94) and slightly lower accuracy in the case of Crossed Eyes data (Accuracy = 93.20%, F1-score = 0.90). The overall accuracy of the MobileNetV2 model across the disease classes reaches 94.85%.

Table 4: Mobile Net V2 Performance Metrics

Class	Accuracy	Precision	Recall	F1-Score
Cataracts	94.70	0.91	0.92	0.92
Glaucoma	93.80	0.90	0.91	0.91
Bulging Eyes	95.00	0.92	0.92	0.92
Crossed Eyes	93.20	0.90	0.91	0.90
Uveitis	96.40	0.93	0.95	0.94
Average Accuracy	94.58			

The results about classification are indicated in the confusion matrix of Figure 9. The highest numbers of correctly classified were achieved in most categories, with only Uveitis depicting a high percentage of correct hits (143 correct predictions). There were also many correct classifications in Cataracts and Bulging Eyes; the misclassification was relatively small. There were some minor overlaps between Glaucoma and Cataracts, and between Crossed Eyes and other classes, which had a slight effect on precision and recall.

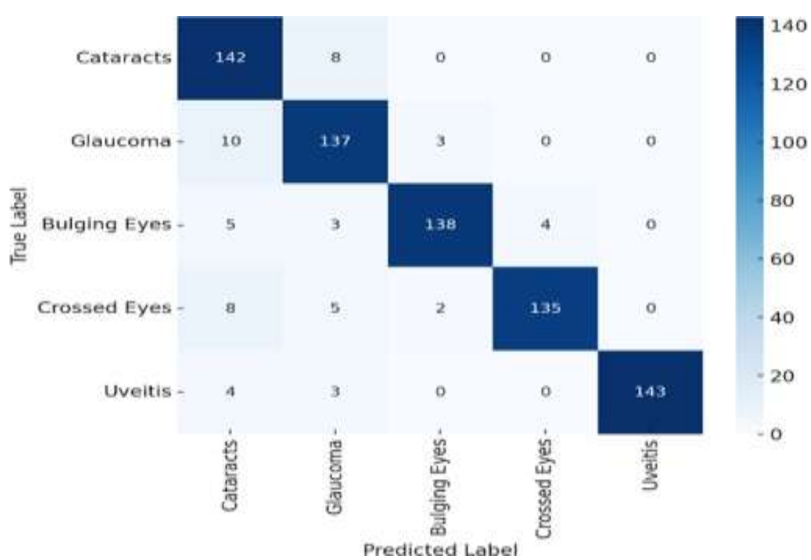

Fig.9: Confusion Matrix of MobileNetV2 Model

Figure 10 displays the learning curve, which indicates how the accuracy of training and validation improves as training progresses through 20 epochs. Both curves have a gradual increasing tendency, and the accuracy of training score is aiming at about 97% accuracy, whereas validation accuracy is scurrying towards about 95%. This implies that the model has strong abilities of generalization to unseen data, and there is a small gap in performance between training and validation. The proximity of both curves also implies low overfitting, and thus, MobileNetV2 is a computationally cost-effective yet robust model at detecting disease.

Fig.10: Learning Curve of MobileNetV2 Model

Ensemble Model Performance Metrics

The Ensemble Model yielded the best performance on classification across all models tested, with a precision rate of 97.93%. The effectiveness of the use of multiple architectures to minimize bias and variance is noticeably reached by this considerable increase in performance compared to single model performance, like that of VGG19 or MobileNetV2.Table 5 demonstrates per-class outcomes, and the results are outstanding in terms of the classification of all categories. Cataracts and Uveitis obtained the two best accuracies of 98.10% and 98.50%, respectively, with precision, recall, and F1-scores properly above 0.95. The overall performance of the ensemble was good, because even in the relatively more challenging class, Crossed Eyes, the accuracy stood at 96.80%.

Table 5: Ensemble Model Performance Metrics

Class	Accuracy	Precision	Recall	F1-Score
Cataracts	98.10	0.98	0.97	0.97
Glaucoma	97.70	0.97	0.95	0.96
Bulging Eyes	97.50	0.97	0.96	0.96
Crossed Eyes	96.80	0.96	0.94	0.95
Uveitis	98.50	0.98	0.97	0.98
Average Accuracy	97.93			

The reliability of the model is also further supported by the confusion matrix of Figure 11, which exhibits very little misclassification. The majority of predictions are aligned along the diagonal plot, with slight overlaps between Glaucoma and Crossed Eyes and between Bulging Eyes and Glaucoma. Such confusions are not surprising as they are due to the overlapping precision in their clinical characteristics; however, they are insignificant to the accuracy of classification.

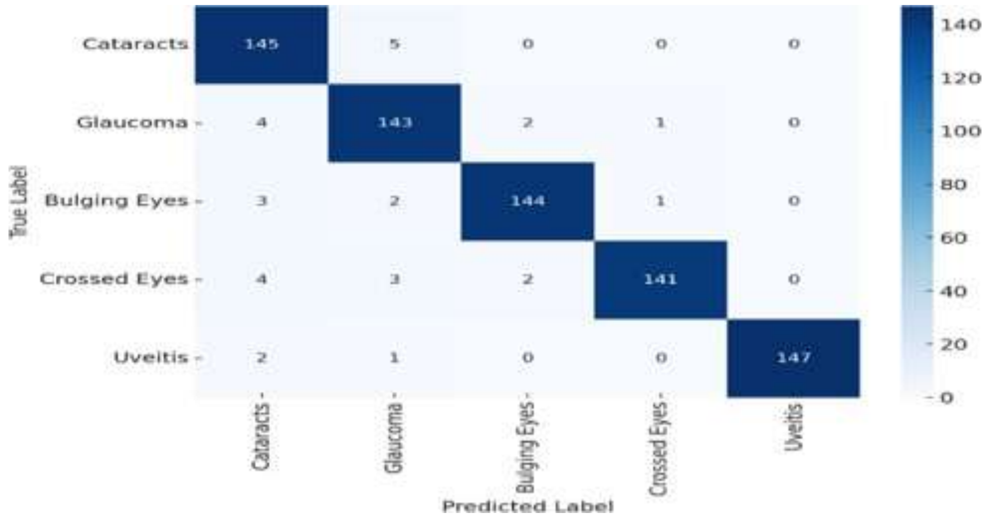

Fig.11: Confusion Matrix of Ensemble Model

In Figure 12, the learning curve shows that training and validation accuracy are gradually increasing across 20 epochs, showing only slight disparity between them. This indicates that the model was generalizable, and it was not overfit. The accuracy of validation is close to the accuracy of training, which implies a good balance between bias and variance.

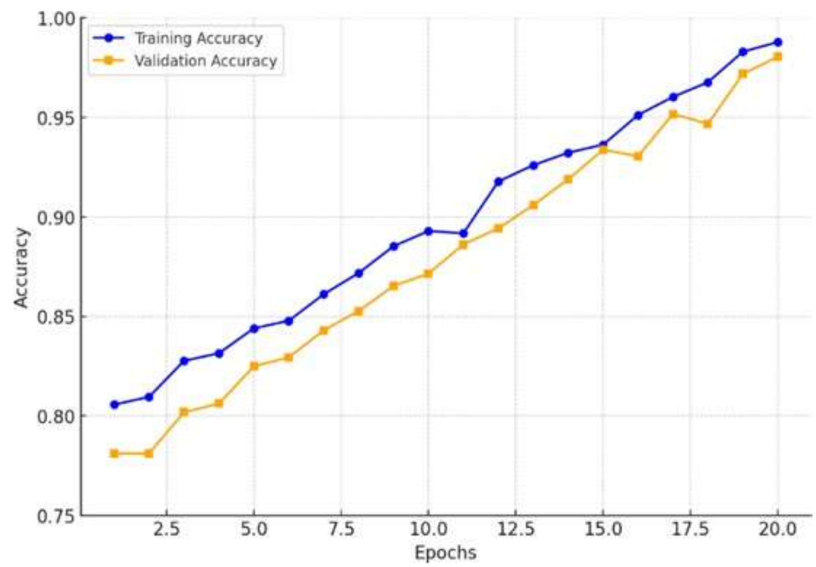

Fig.12: Learning Curve of Ensemble Model

The findings show that the ensemble learning method performed better than individual CNN architectures due to the fact that a combination of the strengths of such architectures was achieved. These results demonstrate the effectiveness of ensembles in the classification of ophthalmic diseases due to the low misclassification rate and higher performance measurements. Specifically, high recall and F1-scores demonstrate that the model is very reliable in making predictions in medical diagnostics, where minimization of false negatives is essential. Among the architectures, the Ensemble Model is found to be the most robust and clinically relevant one, which can give reasonably accurate predictions over all the classes, but avoiding generalization, which becomes an essential feature of the model to be used in real-world ophthalmological applications.

Benchmark Model Comparison

The ensemble was compared with ResNet50, DenseNet121, and EfficientNetB0.As can be seen in Table VI and Figure 13 the ensemble framework demonstrated the best performance of 97.93% classification accuracy, which exceeded prevailing baseline models. Across the individual models, DenseNet121 had the highest accuracy (96.88%), followed by ResNet50 (96.12%) and VGG19(96.07%).The ensemble, however, was usually more accurate, precise, and recalled and achieved a higher F1-score than these models, which is indicative of the advantage that a combination of complementary features between multiple networks can bring.

Table 6: Comparison with Benchmark Models

Model	Accuracy	Precision	Recall	F1-Score
VGG16	95.21	0.94	0.95	0.94
VGG19	96.07	0.95	0.96	0.95
MobileNetV2	94.85	0.94	0.94	0.94
ResNet50	96.12	0.95	0.96	0.95
DenseNet121	96.88	0.96	0.96	0.96
Efficient NetB0	95.75	0.95	0.95	0.95
Ensemble	97.93	0.98	0.98	0.98

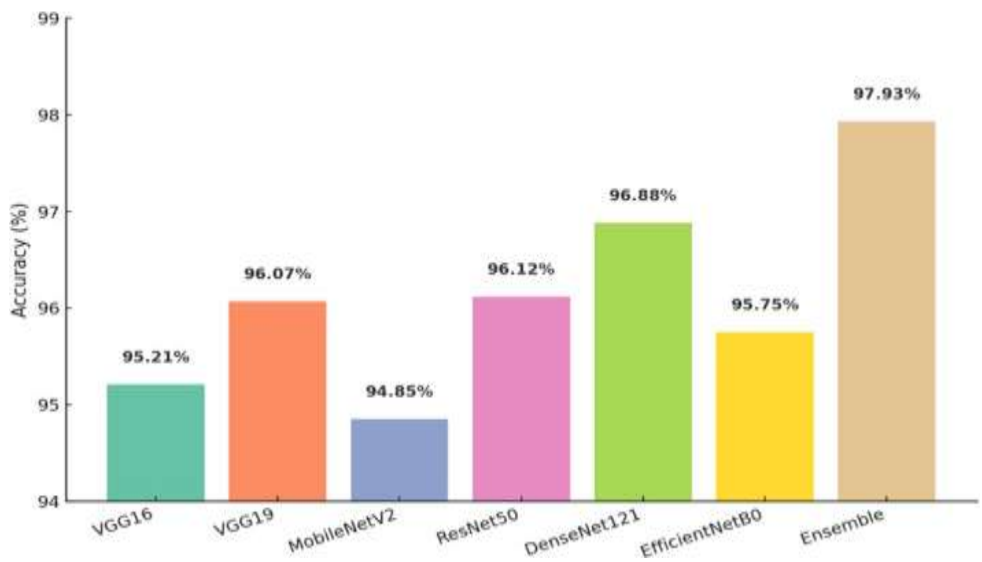

Fig.13: Accuracy Comparison of Ensemble Model with Baseline Models

The results confirm that the ensemble framework outperforms all baseline models in terms of accuracy, precision, recall, and F1-score, while maintaining computational efficiency. Table 7 compares the proposed ensemble deep transfer learning framework with four existing studies, showing superior performance with 97.93% accuracy and 0.98 precision, recall, and F1-Score.Unlikeexistingmodelsthatfocusonsinglediseases or require complex multi-stage processes, the proposed model efficiently handles five eye diseases with high generalizability. This makes it computationally efficient and clinically applicable for real-time use.

Table 7: Comparison of Related Studies with the Proposed Model

Study	Accuracy	Precision	Recall	F1-Score
[41]	95%	90%	93%	90%
[42]	94%	94%	93%	93%
[43]	96.2%	96%	95%	96%
[44]	93%	94%	94%	94%
Proposed (Ensemble)	97.93%	98%	98%	98%

6. Discussion

The experimental findings established that deep transfer learning models are effective in multiclass eye disease recognition. Overall, VGG19 and DenseNet121 were more accurate, achieving 96.07% and 96.88%, respectively, among the individual models. These findings imply that more profound architectures that have better feature extraction abilities are more suitable to dealing with complicated changes in ocular pictures. The MobileNetV2 model, being relatively lightweight and practical, yielded relatively lower accuracy (94.85%), indicating that model compactness can come at the cost of discrimination in medical image classification. One of the most valuable contributions of this work is an ensemble learning framework, which integrates the advantages of the VGG16, VGG19, and MobileNetV2 models. The ensemble model was far superior to individual models with an accuracy of 97.93 percent, precision of 0.98, recall of 0.98, and F1-score of 0.98. This finding validates the hypothesis that in the multiple models framework, the robustness of the classification model can be improved through the utilization of diverse feature representations, thereby cutting back its single architecture constraints. The ensemble learning curves also demonstrated more steady convergence and the least overfitting as compared to standalone models. The proposed ensemble framework also showed improved results when compared to benchmark deep learning models, including ResNet50, DenseNet121, and EfficientNetB0. Although DenseNet121 yielded promising results (accuracy of 96.88%), the ensemble exceeded it by more than 1%, which is an impressive result in terms of medical image classification applications, which require diagnostic accuracy. These results suggest the possibility of using deep transfer learning ensemble strategies in ophthalmological diagnosis to increase patient accuracy. Another significant finding is the performance by classes of different eye diseases. The ensemble model showed a significant sensitivity and specificity to detect Cataracts (98.1%) and Uveitis (98.5%), which are usually difficult to detect given their shared symptoms with other maladies. The recall gains in each class of the disease show that the ensemble minimizes false negatives, which is essential to early detection and treatment planning.

7. Conclusion

This study presented a comprehensive framework for multiclass eye disease recognition and classification using deep transfer learning and an ensemble learning strategy. By employing state-of-the-art pre-trained models such as VGG16, VGG19, and MobileNetV2, and integrating them into an ensemble, the research demonstrated that combining complementary feature extraction capabilities significantly enhances diagnostic performance. The experimental evaluation confirmed that while individual models like VGG19 (96.07% accuracy) and DenseNet121 (96.88% accuracy) provided strong results, the ensemble model outperformed all others, achieving 97.93% accuracy, 0.98 precision, 0.98 recall, and 0.98 F1-score. This superior performance highlights the effectiveness of ensemble learning in handling complex and diverse ophthalmic image data.Notably,the ensemble demonstrated strong class-wise performance, particularly in detecting Cataracts and Uveitis, reducing the risk of false negatives and increasing diagnostic reliability. Furthermore, comparison with benchmark architectures such as ResNet50, DenseNet121, and EfficientNetB0 validated the robustness of the proposed approach. The ensemble framework not only surpassed individual baselines but also provided smoother learning curves and more stable generalization, minimizing overfitting. This study establishes that ensemble deep transfer learning is a promising direction for computer-aided diagnosis (CAD) in ophthalmology by improving accuracy and reliability in multiclass eye disease detection, the proposed framework has the potential to support ophthalmologists in early diagnosis, reduce clinical workload, and ultimately contribute to better patient outcomes. Future research can focus on expanding the dataset with more diverse and clinically validated images to improve generalization. Explain ability techniques such as Grad-CAM should be integrated to enhance model interpretability. Deployment on mobile and edge devices will allow real-time clinical use, especially in low-resource settings. Finally, combining retinal images with patient metadata or other imaging modalities could further improve diagnostic accuracy.

DATA AVAILABILITY STATEMENT

The dataset used in this study is publicly available on Kaggle at: https://www.kaggle.com/datasets/kondwani/eye-disease-dataset. This dataset includes images across five eye disease categories Bulging Eyes, Cataracts, Crossed Eyes, Glaucoma, and Uveitis matching the classification tasks in the manuscript

REFERENCES

[1] T. Babaqi, M. Jaradat, A. E. Yildirim, S. H. Al-Nimer, and D. Won, "Eye disease classification using deep learning techniques," arXiv preprint arXiv: 2307.10501, 2023.

[2] A. M. Gajaram, "An approach to classify ocular diseases using machine learning and deep learning," Ph.D. dissertation, Dublin, National College of Ireland, 2022.

[3] K. B. Vardhan, M. Nidhish, D. Shameem et al., "Eye disease detection using deep learning models with transfer learning techniques," ICST Transactions on Scalable Information Systems, vol. 11, pp. 1–13, 2024.

[4] A. Choudhary, S. Ahlawat, S. Urooj, N. Pathak, A. Lay-Ekuakille, and N. Sharma, "A deep learning-based framework for retinal disease classification," in Healthcare, vol. 11, no. 2. MDPI, 2023, p. 212.

[5] S. Muchuchuti and S. Viriri, "Retinal disease detection using deep learning techniques: a comprehensive review," Journal of Imaging, vol. 9, no. 4, p. 84, 2023.

[6] B. Goutam, M. F. Hashmi, Z. W. Geem, and N. D. Bokde, "A comprehensive review of deep learning strategies in retinal disease diagnosis using fundus images," IEEE Access, vol. 10, pp. 57 796–57 823, 2022.

[7] R. C. Joshi, A. K. Sharma, and M. K. Dutta, "Visiondeep-ai: Deep learning-based retinal blood vessels segmentation and multi-class classification framework for eye diagnosis," Biomedical Signal Processing and Control, vol. 94, p. 106273, 2024.

[8] P. Kaushik and P. Sharma, "Optimizing automated eye disease detection: A deep learning approach using efficientnetb3 for accurate multi-class classification," in 2024 IEEE International Conference on Intelligent Signal Processing and Effective Communication Technologies (INSPECT). IEEE, 2024, pp. 1–6.

[9] S. K. Dash, P. K. Sethy, A. Das, S. Jena, and A. Nanthaamornphong, "Advancements in deep learning for automated diagnosis of ophthalmic diseases: a comprehensive review," IEEE Access, 2024.

[10] A. Bilal, L. Zhu, A. Deng, H. Lu, and N. Wu, "Ai-based automatic detection and classification of diabetic retinopathy using u-net and deep learning," Symmetry, vol. 14, no. 7, p. 1427, 2022.

[11] A. A. Abd El-Khalek, H. M. Balaha, N. S. Alghamdi, M. Ghazal, A. T. Khalil, M. E. A. Abo-Elsoud, and A. El-Baz, "A concentrated machine learning-based classification system for age-related macular degeneration (amd) diagnosis using fundus images," Scientific Reports, vol. 14, no. 1, p. 2434, 2024.

[12] P. Thiagarajan and M. Suguna, "Deep learning ocular disease detection system (odds)," in International Conference on Mining Intelligence and Knowledge Exploration. Springer, 2023, pp. 213–224.

[13] M. Khalid, M. Z. Sajid, A. Youssef, N. A. Khan, M. F. Hamid, and F. Abbas, "Cad-eye: An automated system for multi-eye disease classification using feature fusion with deep learning models and fluorescence imaging for enhanced interpretability," Diagnostics, vol. 14, no. 23, p.

2679, 2024.

[14] A. K. Bhandari, S. Shahnawazuddin et al., "Automated computationally intelligent methods for ocular vessel segmentation and disease detection: A review." Archives of Computationa l Methods in Engineering, vol. 31, no. 2, 2024.

[15] V. Rajyaguru, C. Vithalani, and R. Thanki, "A literature review: various learning techniques and its applications for eye disease identification using retinal images," International Journal of Information Technology, vol. 14, no. 2, pp. 713–724, 2022.20

[16] P. Kaushik and P. Sharma, "Efficient eye disease detection: A deep learning approach with inceptionresnetv2 for retinal image classification," in 2024 International Conference on Advances in Computing, Communication and Materials (ICACCM). IEEE, 2024, pp. 1–5.

[17] F. Du, L. Zhao, H. Luo, Q. Xing, J. Wu, Y. Zhu, W. Xu, W. He, and J. Wu, "Recognition of eye diseases based on deep neural networks for transfer learning and improved ds evidence theory," BMC Medical Imaging, vol. 24, no. 1, p. 19, 2024.

[18] S. Narayanan et al., "Deep learning of fundus images and optical coherence tomography images for ocular disease detection–a review," Multimedia Tools and Applications, vol. 83, no. 41, pp. 88 745–88 789, 2024.

[19] A. A. Soofi et al., "Exploring deep learning techniques for glaucoma detection: a comprehensive review," arXiv preprint arXiv: 2311.01425, 2023.

[20] A. Rahman, M. Youldash, G. Alshammari, A. Sebiany, J. Alzayat, M. Alsayed, M. Alqahtani, and N. Aljishi, "Diabetic retinopathy detection: A hybrid intelligent approach." Computers, Materials & Continua, vol. 80, no. 3, 2024.

[21] S. K. Sharma and D. Muduli, "Automated diagnosis model for glaucoma detection: A deep learning feature fusion and ls-svm based approach," in 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2023, pp. 1–6.

[22] R. Thakur et al., "Optimizing retinal disease detection: A comparative analysis of state-of-the-art deep learning models for multiclass classification," in 2025 International Conference on Networks and Cryptology (NETCRYPT). IEEE, 2025, pp. 638–643.

[23] J. Hernández-Capistran, L. N. Sánchez-Morales, G. Alor-Hernández, M. Bustos-López, and J. L. Sánchez-Cervantes, "Machine and deep learning algorithms for adhd detection: A review," Innovations in machine and deep learning: Case studies and applications, pp. 163–191, 2023.

[24] C. Choudhary, A. Mathur, and R. Gupta, "Diabetic retinopathy detection: A transfer learning based approach for accurate diagnosis," in 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS). IEEE, 2023, pp. 280–285.

[25] G. Nokas and T. Kotsilieris, "Preventing keratoconus through eye rubbing activity detection: a machine learning approach," Electronics, vol. 12, no. 4, p. 1028, 2023.

[26] A. Jones and T. B. Vijayan, "Ai-driven canine cataract detection: a machine learning approach using support vector machine," Journal of the Chinese Institute of Engineers, vol. 48, no. 1, pp. 94–102, 2025.

[27] A. R. Ran, X. Wang, P. P. Chan, M. O. Wong, H. Yuen, N. M. Lam, N. C. Chan, W. W. Yip, A. L. Young, H.-W. Yung et al., "Developing a privacy preserving deep learning model for glaucoma detection: a multicentre study with federated learning," British Journal of Ophthalmology, vol. 108, no. 8, pp. 1114–1123, 2024.

[28] A. Mutawa, S. Alnajdi, and S. Sruthi, "Transfer learning for diabetic retinopathy detection: A study of dataset combination and model performance," Applied Sciences, vol. 13, no. 9, p. 5685, 2023.

[29] G. Zhang, J.-W. Lin, J. Wang, J. Ji, L.-P. Cen, W. Chen, P. Xie, Y. Zheng, Y. Xiong, H. Wu et al., "Automated multidimensional deep learning platform for referable diabetic retinopathy detection: a multicentre, retrospective study," BMJ open, vol. 12, no. 7, p. e060155, 2022.

[30] M. A. Al Jbaar and S. A. Dawwd, "Dcnn-based embedded models for parallel diagnosis of ocular diseases." Eastern-European Journal of Enterprise Technologies, vol. 124, no. 2, 2023.

[31] A. Kaur, V. Kukreja, M. Kumar, A. Choudhary, and R. Sharma, "Retdense: A fine-tuned densenet121 model for retinal eye disease detection," in 2024 IEEE International Conference on

Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), vol. 2. IEEE, 2024, pp. 1–5.
[32] D. Selvathi, "Classification of ocular diseases using transfer learning approaches and glaucoma severity grading," in Computational methods and deep learning for ophthalmology. Elsevier, 2023, pp. 1–15.
[33] D. Kumar and C. Ravindra, "Transfer learning and hybrid models for the classification of eye diseases on retinal images," in Applications of Machine Learning 2024, vol. 13138. SPIE, 2024, pp. 211–221.
[34] M. A. Priyadarshini, S. Salma, D. Sailesh, E. Manasa, G. L. Charan, and B. Dinesh, "A visionary approach to anemia detection: Integrating eye condition data and machine learning," in International Conference on Computational Innovations and Emerging Trends (ICCIET-2024). Atlantis Press, 2024, pp. 781–793.
[35] Z. A. Haider, N. A. Alsadhan, F. M. Khan, W. Al-Azzawi, I. U. Khan, and I. Ullah, "Deep learning-based dual optimization framework for accurate thyroid disease diagnosis using cnn architectures," Mehran University Research Journal of Engineering and Technology, vol. 44, no. 2, pp. 1–12, 2025.
[36] D. Bhulakshmi and D. S. Rajput, "A systematic review on diabetic retinopathy detection and classification based on deep learning techniques using fundus images," PeerJ Computer Science, vol. 10, p. e1947, 2024.
[37] I. Vohra, D. Tripathi, V. Joshi, K. Rana, and R. Chauhan, "Enhancing microaneurysm detection: A deep learning perspective on false positives and optic disc alignment," in 2024 International Conference on Automation and Computation (AUTOCOM). IEEE, 2024, pp. 618–622.
[38] P. Thiagarajan and M. Suguna, "Deep learning ocular disease detection system (odds)," in International Conference on Mining Intelligence and Knowledge Exploration. Springer, 2023, pp. 213–224.
[39] S. Vidivelli, P. Padmakumari, C. Parthiban, A. DharunBalaji, R. Manikandan, and A. H. Gandomi, "Optimising deep learning models for ophthalmological disorder classification," Scientific Reports, vol. 15, no. 1, p. 3115, 2025.
[40] M. Wassel, A. M. Hamdi, N. Adly, and M. Torki, "Vision transformers based classification for glaucomatous eye condition," in 2022 26th International Conference on Pattern Recognition (ICPR). IEEE, 2022, pp. 5082–5088.
[41] S. Gulati, K. Guleria, and N. Goyal, "Detection and multiclass classification of ocular diseases using deep learning-based ensemble model," Int. J. Intell. Syst. Appl. Eng, vol. 12, pp. 18–29, 2024.
[42] S. Al-Fahdawi, A. S. Al-Waisy, D. Q. Zeebaree, R. Qahwaji, H. Natiq, M. A. Mohammed, J. Nedoma, R. Martinek, and M. Deveci, "Fundus-deepnet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images," Information Fusion, vol. 102, p. 102059, 2024.
[43] H. Yu and X. Dong, "Ensemble-based eye disease detection system utilizing fundus and vascular structures," Scientific Reports, vol. 15, no. 1, p. 19298, 2025.
[44] Y.-T. Lin, X. Xiong, Y.-P. Zheng, and Q. Zhou, "Transfer learning and multi-feature fusion-based deep learning model for idiopathic macular hole diagnosis and grading from optical coherence tomography images," Clinical Ophthalmology, pp. 1593–1607, 2025.
[45] M. A. Khan, M. Z. Hussain, M. F. Khan, M. Ahmad, S. Abbas, T. Mazhar, T. Shahzad, and M. M. Saeed, "Blockchain enabled iomt and transfer learning for ocular disease classification," Discover Applied Sciences, vol. 7, no. 5, pp. 1–19, 2025.
[46] S. Vidivelli, P. Padmakumari, C. Parthiban, A. DharunBalaji, R. Manikandan, and A. H. Gandomi, "Optimising deep learning models for ophthalmological disorder classification," Scientific Reports, vol. 15, no. 1, p. 3115, 2025.
[47] M. R. Shoaib, H. M. Emara, J. Zhao, W. El-Shafai, N. F. Soliman, A. S. Mubarak, O. A. Omer, F. E. Abd El-Samie, and H. Esmaiel, "Deep learning innovations in diagnosing diabetic retinopathy: The potential of transfer learning and the diacnn model," Computers in Biology and

Medicine, vol. 169, p. 107834, 2024.
[48] Y. S. Inan, “Adaptive multiscale retinal diagnosis: A hybrid trio-model approach for comprehensive fundus multi-disease detection leveraging transfer learning and siamese networks,” arXiv preprint arXiv: 2405.18449, 2024.
[49] D. Shi, W. Zhang, X. Chen, Y. Liu, J. Yang, S. Huang, Y. C. Tham, Y. Zheng, and M. He, “Eyefound: a multimodal generalist foundation model for ophthalmic imaging,” arXiv preprint arXiv:2405.11338, 2024.
[50] P. Gu, T. Yao, M. He, F. Duan, F. Liu, R. Peng, and B. Ge, “Swinecat: A transformer-based fundus disease classification model with shifted window attention and efficient channel attention,” arXiv preprint arXiv:2507.21922, 2025.
[51] E. S. Kumar and C. S. Bindu, “Mdcf: Multi-disease classification framework on fundus image using ensemble cnn models,” Journal of Jilin University,vol. 40, no. 09, pp. 35–45, 2021
[52] T. Pratap and P. Kokil, “Computer-aided diagnosis of cataract using deep transfer learning,” Biomedical Signal Processing and Control, vol. 53, p. 101533, 2019.
[53] K. B. Vardhan, M. Nidhish, D. Shameem et al., “Eye disease detection using deep learning models with transfer learning techniques,” ICST Transaction on Scalable Information Systems, vol. 11, pp. 1–13, 2024.
[54] G. D. Deepak and S. K. Bhat, “Deep learning-based cnn for multiclassification of ocular diseases using transfer learning,” Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, vol. 12, no. 1, p. 2335959, 2024.
[55] Q. Abbas, M. Albathan, A. Altameem, R. S. Almakki, and A. Hussain, “Deep-ocular: improved transfer learning architecture using self-attention and dense layers for recognition of ocular diseases,” Diagnostics, vol. 13, no. 20, p. 3165, 2023.
[56] M. Z. Muntaqim, T. A. Smrity, A. S. M. Miah, H. M. Kafi, T. Tamanna, F. Al Farid, M. A. Rahim, H. A. Karim, and S. Mansor, “Eye disease detection enhancement using a multi-stage deep learning approach,” IEEE Access, 2024.
[57] M. G. Gualsaqui, S. M. Cuenca, I. L. Rosero, D. A. Almeida, C. Cadena, F. Villalba, and J. D. Cruz, “Multi-class classification approach for retinal diseases,” Journal of Advances in Information Technology, vol. 14, no. 3, pp. 392–398, 2023.
[58] A. Nawaz, T. Ali, G. Mustafa, M. Babar, and B. Qureshi, “Multi-class retinal diseases detection using deep cnn with minimal memory consumption,” IEEe Access, vol. 11, pp. 56 170–56 180, 2023.
[59] M. Ismail, M. Islam, I. Ahmad, F. A. Khan, A. B. Qazi, Z. H. Khan, Z. Wadud, and M. Al-Rakhami, “Reliable path selection and opportunistic routing protocol for underwater wireless sensor networks,” IEEE Access, vol. 8, pp. 100 346–100 364, 2020.
[60] A. Bali and V. Mansotra, “Transfer learning-based one versus rest classifier for multiclass multi-label ophthalmological disease prediction,” International Journal of Advanced Computer Science and Applications, vol. 12, no. 12, 2021.