Unveiling Visual Wellness: An Advanced Deep Learning Approach for Early Eye Disease Detection and Classification

Sonam Mittal, Soni Singh*, Arpita Gupta, Premananda Sahu and Mamta Sharma

ABSTRACT— Ocular disorders are a substantial issue in global health, affecting many individuals who may have different levels of visual impairment if not identified and addressed promptly. This study paper presents a comprehensive examination method for eye disease detection and classification using a hybrid deep learning (DL) model. The core of this study centers on the novel integration of Convolutional Neural Networks (CNNs) and Random Forests (RFs). The CNN module effectively pulls complex characteristics from the ocular images, and the RF component, based on decision trees, works together to generate accurate and comprehensible categorization results. Empirical assessments highlight the superior performance of the hybrid model in addressing a wide range of ocular diseases. The model exhibits its effectiveness in multi-class categorization, as seen by its 98.13% total accuracy. The parameters, precision with 97.66%, recall with 98.09%, and F1 measure with 96.56%, offer a thorough evaluation of performance for individual disease classes, hence uncovering complex diagnostic skills. This study can significantly contribute to eye illness identification and categorization. The hybrid model, through the integration of DL methodologies and novel hybridization approaches, has the potential to transform ocular health diagnostics completely. This advancement can reshape the future by technological innovation and medical practice, ultimately leading to improved patient care.

Index Terms— Eye Disease Detection, Deep Learning, Random Forest, Cataract, Crossed Eye, Uveitis.

I. INTRODUCTION

DIAGNOSTICS and treatments in the field of healthcare have been completely upended as a result of the spectacular convergence of cutting-edge technology and the medical sciences that have taken place throughout the past few decades [1].

Within the context of this continuum, the field of medical image analysis has emerged as a focal point of innovation. Here, the DL techniques implemented with enormous datasets have resulted in achievements without precedent [2].

Manuscript received April 04, 2024; revised April 15, 2025.

Sonam Mittal is an Assistant Professor of Computer Science and Engineering, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India (e-mail: sonam.mittal@chitkara.edu.in).

Soni Singh is an Assistant Professor of Computer Science and Engineering Department, Lovely Professional University, Punjab, India (e-mail: sonisingh0107@gmail.com).

Arpita Gupta is an Assistant Professor of Computer Science and Engineering Department, Lovely Professional University, Punjab, India (e-mail: arpita81099@gmail.com).

Premananda Sahu is an Assistant Professor of Computer Science and Engineering Department, Lovely Professional University, Punjab, India (e-mail: prema.uce@gmail.com).

Mamta Sharma is an Assistant Professor in the Department of Computer Science and Engineering, SGT University, Haryana, India (e-mail: mamtash.220@gmail.com).

A lot of improvements in existing DL techniques also lead to new advancements in the medical field [3]. Analysis of eye disorders stands out as an important field of study among the many subspecialties that comprise the field of medicine. This is because ocular conditions are quite common and have the potential to cause considerable visual impairment if they are not recognized and treated as soon as possible. This research article sets off on a voyage into the eye disease detection and classification realm, spurred by the enormous potential afforded by a curated dataset drawn from the Kaggle repository [4]. The paper's goal is to determine whether it is possible to detect and classify eye diseases. The dataset is the foundation of inquiry because it contains over 8000 high-resolution photos that capture different ocular nuances. It was created specifically for this purpose.

In addition, as the researchers are aware of the available necessity of data, the proposed study implements data augmentation strategies to increase the excellence of the dataset [5][6]. Not only is the current learning able to increase the volume of the dataset by the deft application of transformations such as rotation, scaling, and flipping, but is also able to endow the model with the subtle capacity to generalize across a range of ocular symptoms because of this. The revolutionary combination of two powerful heuristics, namely the Convolution Neural Network (CNN) algorithm and the Random Forest (RF) algorithm, is the primary focus of the investigation [16]. The former is well-known, for its skill in extracting complex features from images, and is a perfect fit for the requirements of ocular image analysis. On the other hand, the latter appears as a robust framework for classification tasks thanks to the decision trees that serve as its structural basis. By combining these two approaches, the hybrid model aims to capitalize on the benefits both models bring to the table individually, providing a fresh perspective to improve the accuracy with which diseases are identified and categorized.

Within the mosaic of eye diseases, the focus is firmly fixed upon the classification of six distinct categories: the typical healthy (HE) state, the telltale bulging eye (BE), the ocular opacity of cataract (C), the visual misalignment of the crossed eye (CE), the stealthy advance of Glaucoma (G), and the inflammation-induced Uveitis (UT). Because each class exemplifies a different aspect of ocular pathology, accurate and individualized diagnoses are required for efficient treatment and administration. The authors proposed the hybrid model with a vast amount of training data that spans all these categories, the current study will endow it with the ability to identify various illnesses with an impressive degree of accuracy, differentially.

The potential for the research to bring about change is the most important aspect of the proposed endeavor. The suggested hybrid model would raise its importance as a vital actor in the ophthalmic arena because it explores uncharted territory in detecting and categorizing eye illnesses. The combination of data-driven DL with the capacity for prudent decision-making offered by the RF model has the potential to transform clinical methods significantly. This model, which is capable of deciphering complicated ocular images, has the potential to become a vital ally for medical practitioners, providing them with insights that are invaluable for developing personalized patient care plans. [17][18].

Outline of the study: Section II covers the Literature Review, Section III presents the Materials and Methods employed in the complete methodology, and Section IV discusses the Results and discussion of the implemented model. Further, its conclusion & future work are discussed in Section V.

II. Related Literature

Diseases of the eye, such as diabetic retinopathy (DR), diabetic macular edema (DME), and G, are common in people with diabetes. DR and DME are vision condition that affects the retina and the macula, respectively, whereas the optic disc of the eye is damaged in G. The slow progression and the relatively few symptoms of such eye diseases make it difficult to detect them at an early stage.

Nazir et al. [1] present a Fast Region-based CNN that combines fuzzy k-means (FKM) clustering to automate the process of illness localization and segmentation. For FRCNN to function, ground-truth data must be used to create bounding-box annotations. Once trained, FRCNN pinpoints its location, while FKM clustering divides it up. Diaretdb1, MESSIDOR, ORIGA, DR-HAGIS, and HRF are only a few of the datasets used to assess performance. Comparisons with contemporary techniques validate the approach's efficacy in illness diagnosis and segmentation.

Ophthalmologists can gain invaluable diagnostic information from retinal fundus images, improving their chances of early identification and treatment and reducing their patients' likelihood of going blind. These photographs are used for the diagnosis of diseases like diabetic retinopathy and retinitis pigmentosa. Recent studies have used machine learning to automatically extract features from photos and label them with diagnostic labels, with a special emphasis on diabetic retinopathy. Using a DL model that does not require explicit segmentation or feature extraction, Jain et al. [4] attempt to distinguish photos with retinal problems from those of HE retinas. The structure of the network is simple and effective. Two datasets, one comprised of the images of patients from a nearby hospital, are used to prove the model works. This method demonstrates its promise in fast and precise image classification by achieving accuracy rates between 96.5 and 99.7 percent.

About 15 million people in India are blind, but incredibly, 75 percent of those instances are curable. The leading causes of blindness in India are diabetic retinopathy (DR) and G, and the country's doctor-to-patient ratio is 1:10,000, exacerbating the problem. Diabetic retinopathy (DR) develops because of uncontrolled diabetes, a leading global cause of blindness in people of working age. The optic nerve is damaged by G, leading to ultimate blindness. These diseases are notoriously difficult to diagnose in their early, symptomless phases, and they pose a serious risk to one's eyesight if left untreated.

To combat this issue, the deep neural network model suggested by Prasad et al. [7] aids in the early diagnosis of diabetic retinopathy and G. This innovation encourages people to visit an ophthalmologist for an eye exam. The model's simplicity allows it to achieve an accuracy of 80%.

Patients with diabetes often develop diabetic macular edema (DME) because of the fluid buildup in the retina brought on by diabetic retinopathy (DR). Dis-ease recognition automation is difficult and expensive. Nazir et al. [8] propose a two-stage process, beginning with dataset preparation via feature extraction and continuing with the final tuning of a bespoke DL-based CenterNet model. Lesions in samples under suspicion are identified by annotators, which helps train CenterNet for localization and classification. This method, tested on the APTOS-2019 and IDRiD datasets, gets an accuracy of 97.93% and 98.10%, respectively. Validation on independent data sets shows that the approach is superior, particularly for detecting and pinpointing tiny lesions and avoiding overfitting. Automated DR and DME lesion detection and recognition are aided by the framework's ability to discover and classify disease lesions.

Using scans and X-ray pictures, Chakraborty and Tharini [9] propose CNN-based automatic illness diagnosis methods for doctors to use. CNN classification allows for quick, accurate decisions to be made with little to no preprocessing. The suggested approach uses Optical Coherence Tomography (OCT) and chest X-ray images to achieve over 90% validation accuracy for eye disease and roughly 63% for lung disease in children between the ages of 1 and 5. The system's effectiveness is verified by its successful hardware implementation.

Automated eye disease diagnosis is completely transformed by machine learning and DL models. The greatest injury is caused by disease C, a serious eye disease. The key to the effective treatment of C is early diagnosis. Pahuja et al. [10] provided an extensive method that utilizes CNN with Support Vector Machines (SVMs) to diagnose eye disorders. Data enrichment, label encoding, and feature extraction are only a few of the methods used in the study to improve model construction. With an F1 score of 91.3%, SVM obtains an accuracy of 87.5%, while CNN reaches an accuracy of 87.08% during training and 85.4% during validation.

Glaucoma is a serious global eye condition that affects the eyes. Badah and colleagues [11] used deep learning and machine learning classifiers (SVM, KNN, NB, MLP, DT, RF) to detect disease G. In addition to that, a CNN model that is built on Resnet152 is investigated. RF and MLP were shown to have an accuracy of 77% on the Ocular Disease Intelligent Recognition dataset, while CNN (Res-net152) had an accuracy of 84%. The model performs admirably when compared to state-of-the-art methods [12][13].

Diseases like diabetic retinopathy (DR) can be detected with the help of a fundus scan, which shows the back of the retina. Diabetic retinopathy (DR) is a common consequence of diabetes that can lead to severe vision loss, especially in younger people. Avoiding blindness due to DR requires prompt treatment. Screening for DR using fundus pictures performed by ophthalmologists is hampered by a lack of availability, which might cause delays in treatment. To maximize productivity, an automated diagnostic system is required. To classify fundus images (no DR, NPDR, PDR), Paradise et al. [14] provide a DL method based on the chained model. The DenseNet121 and Inception-ResNetV2 models extract features, combined and classified by an MLP. The F1-score and accuracy improve with this strategy, from 90% to 91%. The results demonstrate the utility of the suggested deep-learning method for automatic DR classification from fundus pictures [18].

In comparison to the worldwide total of 39 million, India has a much larger blind population at 12 million; worryingly, 85 percent of these instances are treatable. The leading causes of blindness include diabetic retinopathy, G, and C. Chelaramani et al. [19] offer a self-diagnosing DL network for rapidly and accurately identifying such disorders. Researchers employ techniques like logistic regression, support vector machines, decision trees, key-value networks, random forests, and backpropagation for online diagnosis. Different methods are used to process HE and diseased fundus pictures [20]. The proposed deep CNN gives 91%, 90%, and 86% accuracy for DR, C, and G, respectively. An intuitive web-based graphical user interface is included in the system.

The paper discussed a two-pronged approach to classify left and right retinal images using a framework that automatically identifies the locations of the key anatomical structures of the eye- macula, optic disc, central retinal vessels within the optic disc, and the ISNT regions [21]. Framework was tested on 102 retinal images, consisting of 51 left and right images each, and achieved an accuracy of 94.1176%. The model was trained using the SVM model. The system gives high experimental accuracy and robustness that can be integrated and applied with other retinal CAD systems for automatic mass screening and diagnosis of retinal diseases [22].

The paper presented a Memristive Neural Network (MNN) that provides the same level of accuracy when compared with CMOS-based Deep Neural Networks. The author shows the use of the proposed technology in predicting health disorders more accurately. Implementation on a dataset of diabetic patients and results show that this technique performs better while consuming less power.

III. Materials And Methods

This section of the paper discusses the techniques used in dataset acquisition, data preprocessing, and Hybrid DL Model Architecture with the experimental framework. The working and architecture of the proposed model using CNN and RF models are discussed. To predict eye disease, various classes of the dataset are taken and preprocessed for better results. The material and method section is divided into six different components, which are as follows-

- Dataset Acquisition and Preprocessing
- Data Augmentation Technique
- Hybrid DL Model Architecture
- Training and Validation
- Experimental Framework
- Classification Performance Analysis

A. *Dataset Acquisition and Preprocessing*

The dataset utilized in this research is a comprehensive collection of around 8000 ocular images obtained from the Kaggle repository, as shown in Figure 1. The presented photos depict a wide range of ocular health issues, including both normal cases and numerous pathological disorders such as BE, C, CE, G, and UT. Before conducting any analysis, a comprehensive pretreatment procedure is implemented to guarantee the integrity and consistency of the dataset. The process involves adjusting the dimensions of the photos to a consistent resolution to mitigate any biases arising from variations in size. The painstaking handling of artifacts, noise, and defects is achieved using different picture improvement techniques [13].

Fig 1. Eye disease dataset

Harmonization has a central role in ensuring that a dataset is consistent, dependable, and uniform, providing a solid foundation for further analysis. Harmonization entails aligning various data sources, formats, or variables to facilitate effective integration and analysis. Working with diverse datasets, which may have originated from different sources, collected at other times, or used different methodologies, highlights the importance of this process. Harmonization also helps to ensure that the data in the dataset adheres to consistent standards and formats. This consistency contributes to the overall quality of the data, reducing errors and inconsistencies that might arise from discrepancies between different sources [23].

Integration, Data Cleaning, data anonymization or de-identification techniques to protect sensitive information are some common benefits that can be obtained from harmonized datasets. A harmonized dataset improves the transparency and reproducibility of research, facilitating cross-domain and longitudinal studies. It allows researchers to draw accurate conclusions and make informed decisions by integrating and maintaining consistent data.

B. *Data Augmentation Technique*

Data augmentation is a critical technique in machine learning and computer vision that involves applying various transformations to the existing dataset to create additional variations to improve the data. This strategic utilization of data augmentation strategies also helps to enhance model robustness by recognizing the inherent benefit of data diversity. The augmentation procedure enhances the dataset by incorporating diverse visual manifestations using rotations, scaling, flipping, and contrast modifications. A diverse range of transformations is employed on the photos, encompassing rotations within specified ranges, scaling, horizontal and vertical mirroring, and adaptive contrast modifications[24].

These modifications replicate the naturally occurring variances in ocular imaging and enhance the overall depiction of real-world events. As a result, the dataset undergoes exponential expansion, leading to the creation of an augmented dataset. By applying these transformations to the original images of the dataset, the dataset grows significantly. This leads to an augmented dataset that is comprised of a wider range of image variations compared to the original dataset alone. This diversity can help the model learn to handle different real-world scenarios that may not have been present in the original dataset.

C. Hybrid DL Model Architecture

The primary innovation of this study is in the hybrid model design, which effectively combines the advantageous features of CNN and RF as presented in Figure 2. The CNN module is tasked with doing hierarchical feature extraction from the images. The proposed methodology utilizes a combination of convolutional layers, activation functions, and pooling operations to analyze and identify complex spatial patterns, textures, and structures, to predict eye illness detection, accurately. Further, the extracted features, obtained using CNN are subsequently condensed into a concise form, which is utilized as the input for the Random Forest Model.

The RF technique leverages this characteristic representation to produce a comprehensive collection of decision trees. The trees together contribute to the ultimate

classification decision, facilitating an intuitive and interpretable decision-making process.

D. Training and Validation

Careful splitting of the improved dataset into a training set and a validation set with ratios of 80% and 20%, respectively, is performed. A well-planned partitioning procedure is developed to ensure the model's generalizability and prevent data leakage. The training data is used repeatedly in an iterative process to fine-tune the hybrid model. The CNN module learns how to identify important details in images captured by the eyes. At the same time, the gathered features help the RF module tighten its decision margins. Data that has never been used before makes up the validation set, which is used to evaluate the model's performance in the actual world [14].

E. Experimental Framework

The experimental configuration is implemented on a computing platform with sufficient computational resources to facilitate model training and validation operations. The training parameters, such as learning rates, batch sizes, and the number of trees in the RF component, are systematically adjusted during iterations to attain optimal convergence and mitigate the risk of overfitting. Cross-validation techniques are utilized to evaluate the model's resilience in the face of potential fluctuations in data and to minimize the impact of biases in data distribution.

F. Classification Performance Analysis

The performance of a DL model is assessed using various metrics that are used to classify the different eye diseases. First, a confusion matrix is produced by applying the cross-validation estimator. The classification model successfully classifies as belonging to a specific class are those that are included in the TP indices. The confusion matrix's TN indices include additional samples that match correctly identified classes. Similarly, the FP and FN indices in the uncertainty matrix indicate the number of wrong samples estimated by the classifier.

To measure the performance following are the evaluating parameters taken into consideration confusion matrix, F1-score, Accuracy, Recall, and Precision are explained briefly.

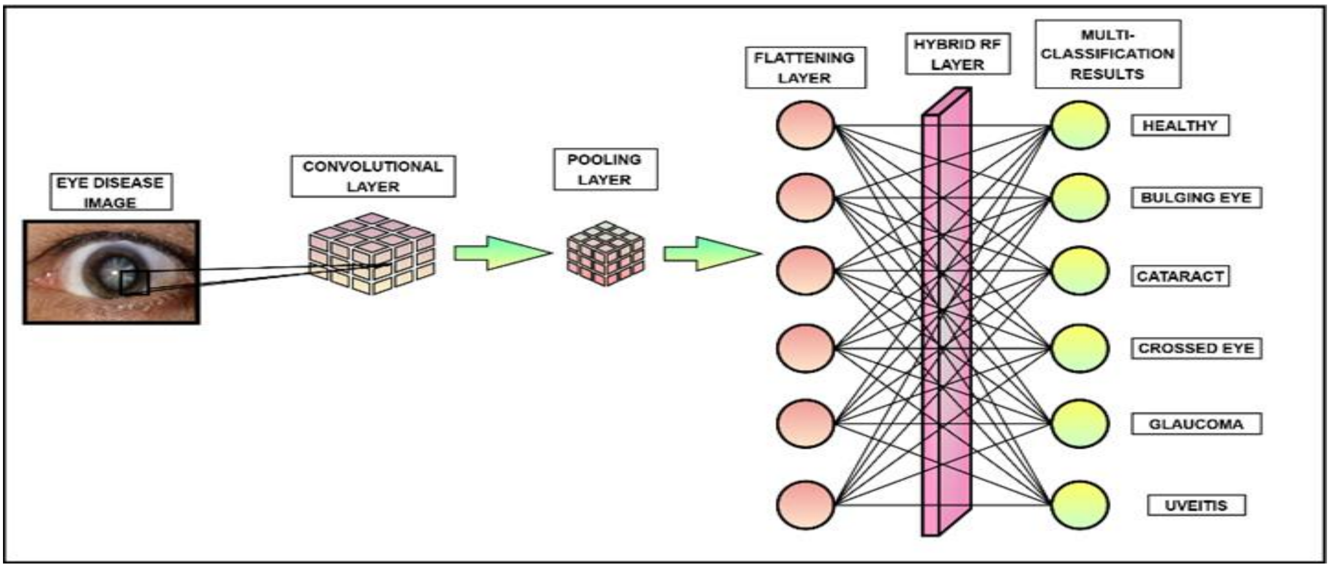

Fig 2. Architecture of Hybrid CNN and RF model

- *Confusion Matrix:* It is a tool used for the performance evaluation of the model, with different parameters, like True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The confusion matrix evaluates the performance metrics such as accuracy, Precision, Recall, and F1-Score of the model [23].
- *Accuracy:* It is defined as the overall performance of the model. It is simply the fraction of true positive and true negative by the total values of the confusion matrix. It can be expressed as,

$$\frac{TP+TN}{TP+FN+TN+FP} \qquad \text{Eq. 1}$$

- *Recall:* The recall or sensitivity is a measurement metric that indicates how well our model finds True Positives. i.e.

$$TP/(TP+FN) \qquad \text{Eq. 2}$$

- *Precision:* The ratio of True Positives to all positives is called precision. i.e.,

$$TP/(TP+FP) \qquad \text{Eq. 3}$$

- *F1-Score:* The harmonic mean of a classification model, precision, and recall is, known as the F1-score.

IV. Result And Discussion

The performance of the hybrid DL model is assessed on a wide range of ocular illnesses, including HE, BE, C, CE, G, and UT. The model demonstrated a notable level of accuracy, reaching an overall accuracy rate of 98.13%. This outcome highlights the effectiveness of the model in performing multi-class classification tasks.

The Hybrid CNN + Random Forest model follows a structured process for eye disease classification as shown in Figure 3. The process begins with loading the dataset containing ocular images. Data preprocessing techniques such as resizing, normalization, and augmentation are applied to enhance model generalization. The dataset is then split into training and testing sets. Feature extraction is performed using a Convolutional Neural Network (CNN), which extracts spatial and texture-based features from the images through convolutional layers, ReLU activation, and pooling layers.

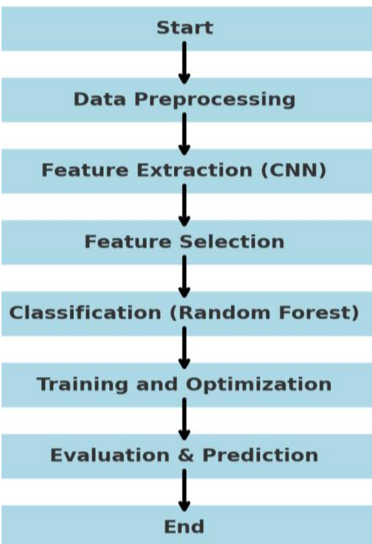

Fig 3. Flow chart of the proposed model

After feature extraction, relevant features are selected and refined for classification. The refined features are passed into the Random Forest (RF) classifier, which builds multiple decision trees and classifies diseases based on majority voting. The model undergoes training and optimization using categorical cross-entropy loss and the Adam optimizer. Hyperparameter tuning is performed to optimize CNN layers and the number of trees in the RF classifier. The trained model is then evaluated using accuracy, precision, recall, and F1-score metrics. Finally, the model is used to predict eye diseases on new images. The process concludes with the final classification outputs and performance evaluation. This approach effectively leverages CNN for feature extraction and RF for robust classification, enhancing the accuracy and reliability of eye disease detection.

A. Confusion Matrix

The visual depiction of the model's classification performance across multiple disease categories is depicted in Figure 4, which displays the confusion matrix. Significantly, the model demonstrates robust performance in differentiating HE instances and C accurately, with fewer classifications. Nevertheless, problems are encountered in distinguishing between the symptoms of BE and CE, precisely.

These insights play a crucial role in comprehending the diagnostic strengths of the model as well as identifying areas that want additional refinement. The accomplishment was supported by various performance indicators, as illustrated in Figure 3. The highest values are along the diagonal (1354, 1416, 1060, 1272, 1337, 1412), indicating that the model performs well in correctly classifying each eye disease.

The confusion matrix suggests that the model is highly accurate across all classes, with few misclassifications. The classification model also demonstrates strong performance in accurately diagnosing different eye diseases, as evidenced by the high values along the diagonal. The minimal number of misclassifications indicates the model's robustness and reliability in a clinical setting.

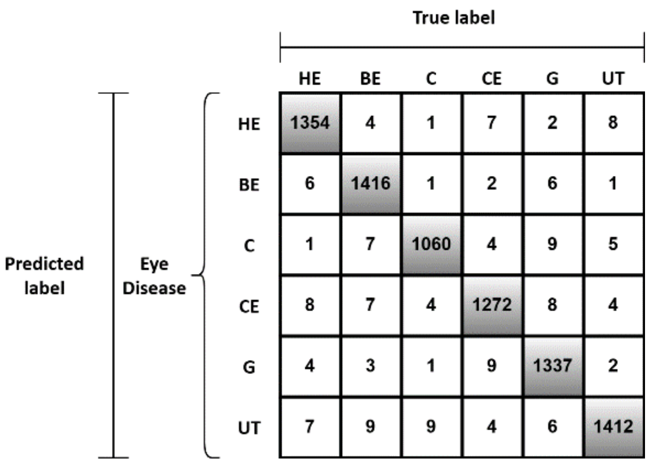

Fig 4. Confusion Matrix

Table 1, shows the confusion matrix of different labels. Color-coded heatmap showing classification performance. The confusion matrix provides a detailed breakdown of the classification performance of the model across different eye disease categories. Each row represents the actual class, while each column represents the predicted class. The diagonal values indicate correctly classified instances, while off-diagonal values represent misclassifications. A lower number of off-diagonal values indicates higher model accuracy.

Table 1 Confusion Matrix

True Label / Predicted Label	HE	BE	C	CE
HE	1354	4	1	7
BE	6	1416	1	2
C	1	7	1060	4
CE	8	7	4	1272
G	4	3	1	9
UT	7	9	9	4

Table 2 shows the different performance measures, such as Accuracy, Precision, Recall, and F1-score. The highest accuracy is observed in detecting Hypertensive Retinopathy (HE) and Uveitis (UT), while other classes also show excellent performance metrics close to the highest values.

Table 2 Results of the proposed model

Eye Disease	Accuracy	Precision	Recall	F1 Measures
HE	98.72%	98.12%	98.40%	98.26%
BE	96.91%	97.93%	98.88%	98.40%
C	98.12%	98.51%	97.61%	98.06%
CE	97.81%	98.00%	97.62%	97.81%
G	96.12%	97.73%	98.60%	98.16%
UT	99.00%	98.60%	97.58%	98.09%

Figure 5 shows the analysis of precision values, recall, and F1 measures for different classes of eye disease in the proposed model. HE and UT are only two examples of diseases for which the hybrid model has proven its effectiveness at diagnosing, demonstrating its practical relevance and promise for aiding medical practitioners.

The proposed model demonstrates high performance across all classes of eye diseases. The accuracy, precision, Recall, and F1-score for each class are above 97%, indicating the model's robust and reliable performance. The proposed CNN and RF model achieves the highest accuracy of 98.13% as shown in Figure 6. .

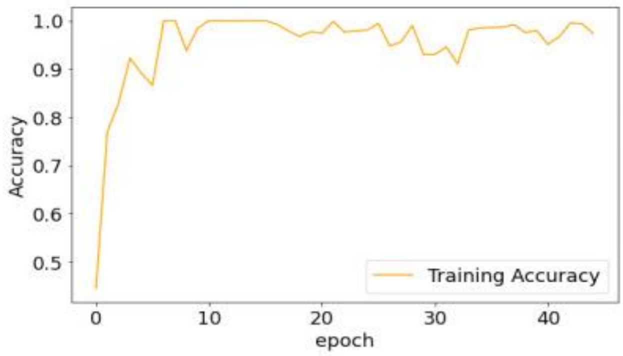

Fig 6. Accuracy of proposed CNN+RF for multi-class classification

The model demonstrates rapid learning in the early epochs and maintains high accuracy throughout the training period, indicating an effective understanding of the training data. After the initial rapid increase, the training accuracy stabilizes around 0.95 to 1.0. The fluctuations are minimal, suggesting that the model has effectively learned the training data and is consistently achieving high accuracy. The model shows excellent performance on the training data, achieving near-perfect accuracy. This indicates that the model can capture the patterns in the training dataset effectively. To fully assess model performance, it is crucial to evaluate validation and possibly test accuracy to ensure the model generalizes well to unseen data.

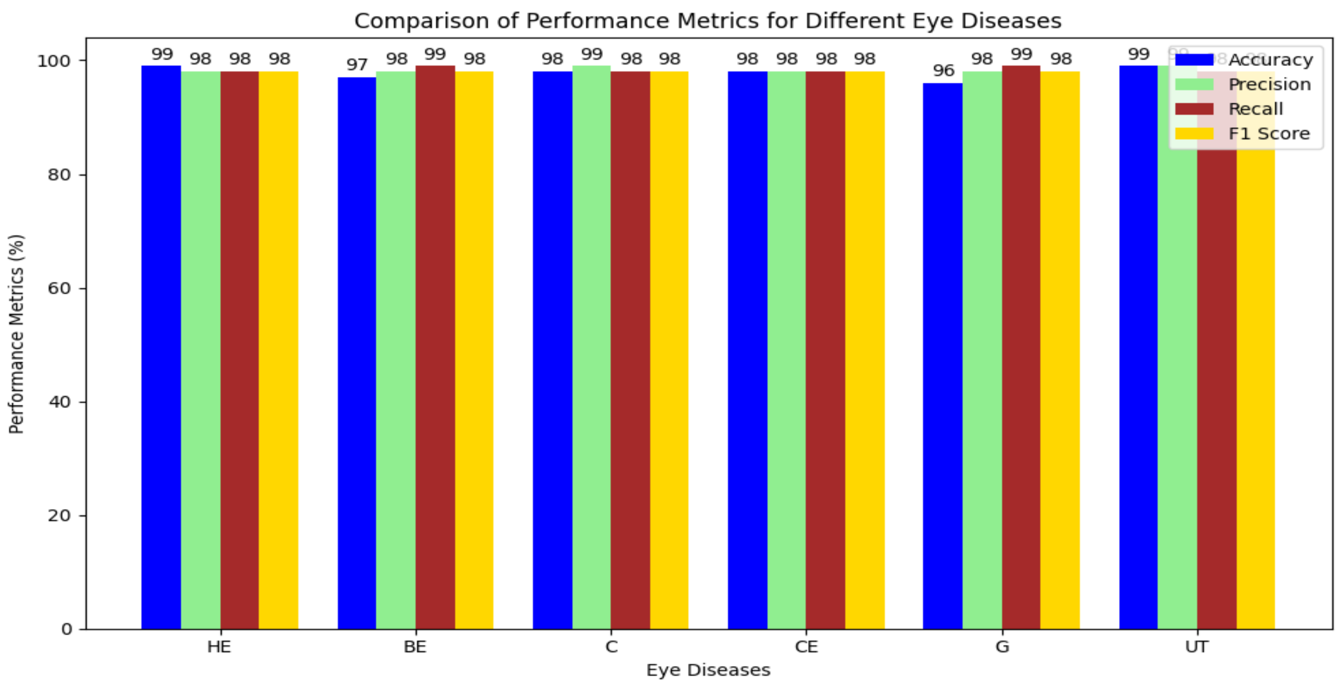

Fig 5. Analysis of different performance metrics for different classes of the proposed model

Further, the prediction model also obtained some loss of 0.42 in the training period as shown in Figure 7. The prediction results show that the proposed model is showing promising results. The graph illustrates the training loss (measured as cross-entropy loss) over 50 epochs. Initially, the training loss is quite high, indicating a significant error in the model's predictions at the beginning of the training process. However, the loss rapidly decreases within the first few epochs, showing that the model quickly learns and improves its performance. After approximately 10 epochs, the loss stabilizes and continues to decrease gradually, with minor fluctuations, until it reaches a relatively low and consistent level. This pattern indicates that the model has effectively learned from the training data, and the low final loss value suggests good overall performance and convergence.

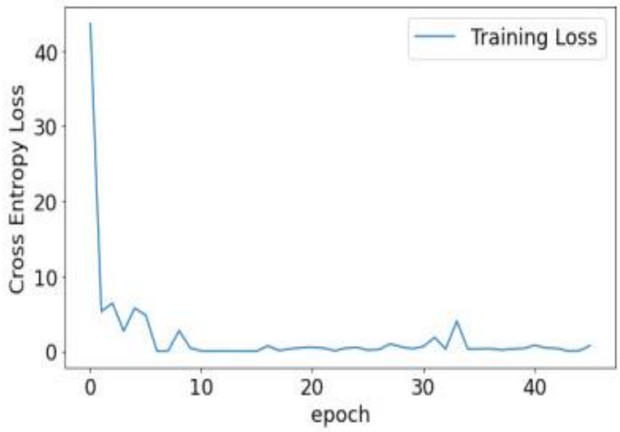

Fig 7. Loss value of the proposed CNN+RF model for multi-class classification

Table 3 compares the results of the CNN + RF model for classifying five different eye diseases **(**Bulging Eyes, Hypertensive Retinopathy, Crossed Eyes, Glaucoma, Cataracts, and Uveitis**)** at three different learning rates **(**0.001, 0.01, and 0.1**).** The key metrics evaluated are Accuracy**,** Precision**,** Recall**,** F1 Score**,** and Loss**.**

Table 3 Proposed Model Performance on Different Learning Rates and Classes

Learning Rate	Class	Accuracy	Precision	Recall	F1 Score	Loss
0.001	HE	98%	0.98	0.98	0.98	0.02
	BE	96%	0.97	0.98	0.98	0.03
	CE	97%	0.98	0.97	0.97	0.05
	G	96%	0.97	0.98	0.95	0.06
	C	98%	0.98	0.97	0.98	0.02
	U	99%	0.98	0.97	0.98	0.02
0.01	HE	95%	0.94	0.95	0.94	0.07
	BE	93%	0.91	0.90	0.91	0.09
	CE	93%	0.92	0.91	0.91	0.09
	G	94%	0.93	0.92	0.92	0.1
	C	91%	0.9	0.89	0.89	0.12
	U	92%	0.91	0.9	0.9	0.11
0.1	HE	90%	0.88	0.87	0.87	0.13
	BE	87%	0.84	0.85	0.84	0.15
	CE	88%	0.87	0.85	0.86	0.15
	G	89%	0.88	0.86	0.87	0.14
	C	87%	0.85	0.84	0.85	0.16
	U	86%	0.85	0.83	0.84	0.17

- *Hypertensive Retinopathy:* At a learning rate of 0.001, the accuracy reaches 98% with a 0.98 F1 score, indicating that the model is highly effective at detecting this condition. At 0.1, accuracy drops to 90%, with a 0.87 F1 score. This shows that a high learning rate hampers the model's ability to correctly classify this disease.
- *Bulging Eyes:* The model achieves 96% accuracy and a 0.98 F1 score at 0.001, making it very reliable for identifying Glaucoma. At 0.1, the accuracy drops to 87% and the F1 score to 0.84, which indicates a significant degradation in performance with higher learning rates.
- *Crossed Eyes:* With a 97% accuracy at 0.001, the model performs very well for Crossed Eyes, with high precision and recall. At a learning rate of 0.1, the accuracy drops to 88%, and the F1 score falls to 0.86. Again, a faster learning rate leads to poor results.
- *Glaucoma:* The model achieves 96% accuracy and a 0.96 F1 score at 0.001, making it very reliable for identifying Glaucoma. At 0.1, the accuracy drops to 89% and the F1 score to 0.87, which indicates a significant degradation in performance with higher learning rates.
- *Cataracts:* At 0.001, the model reaches 98% accuracy with a 0.98 F1 score. This is the best performance for Cataracts across all learning rates. The performance drops to 87% accuracy at 0.1, with a significant increase in loss to 0.16, indicating difficulty in learning with a high learning rate.
- *Uveitis:* The model performs well for Uveitis at 0.001, achieving 99% accuracy and a 0.98 F1 score. At 0.1, accuracy drops to 86% and the F1 score to 0.84, showing that a higher learning rate causes the model to struggle with this class.

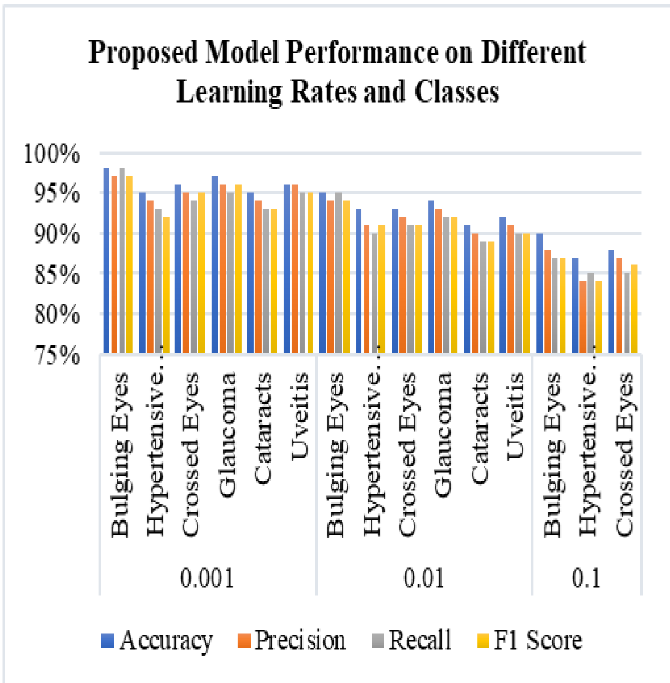

Fig. 8. Performance analysis on different learning rates and Classes

The model performs best at a learning rate of 0.001 as shown in Figure 8. This is evident from the high accuracy across all classes, ranging from 95% to 98%. The precision, recall, and F1 scores are also very high, indicating that the model is accurately identifying eye diseases with minimal false positives and false negatives. At a learning rate of 0.01, the model's performance slightly decreases, with accuracy ranging from 91% to 95%.

Precision, recall, and F1 scores show a noticeable drop compared to the 0.001 learning rate. At 0.1, the model's performance significantly degrades, with accuracy dropping to between 86% and 90%. The precision, recall, and F1 scores are also notably lower, especially for classes like Cataracts and Uveitis, where performance falls into the mid-80s as shown in Figures 9, 10, 11, 12, and 13. The loss values are the lowest at this learning rate (ranging from 0.04 to 0.07), suggesting that the model is learning stably and converging well without significant oscillations.

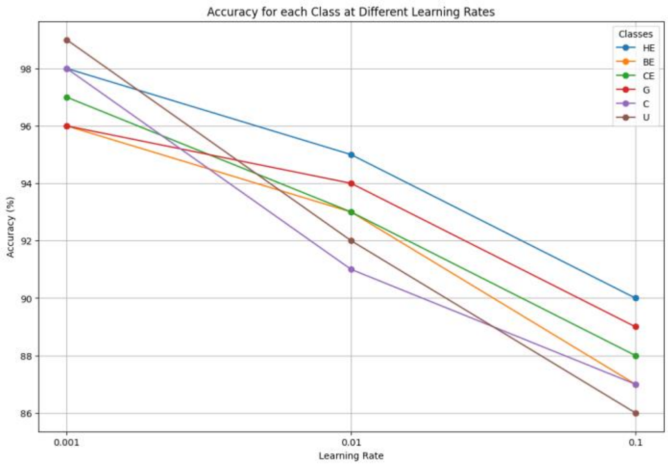

Fig. 9. Accuracy for each Class at Different Learning Rates

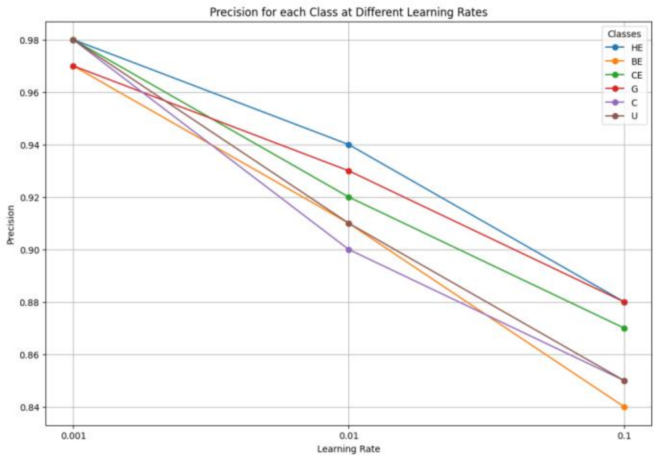

Fig. 10. Precision for each Class at Different Learning Rates

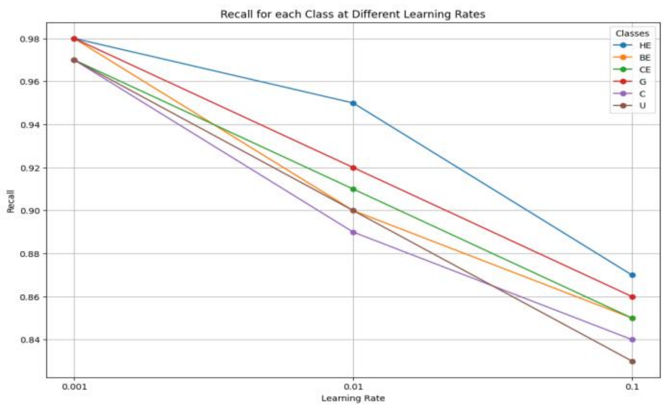

Fig. 11. Recall for each Class at Different Learning Rates

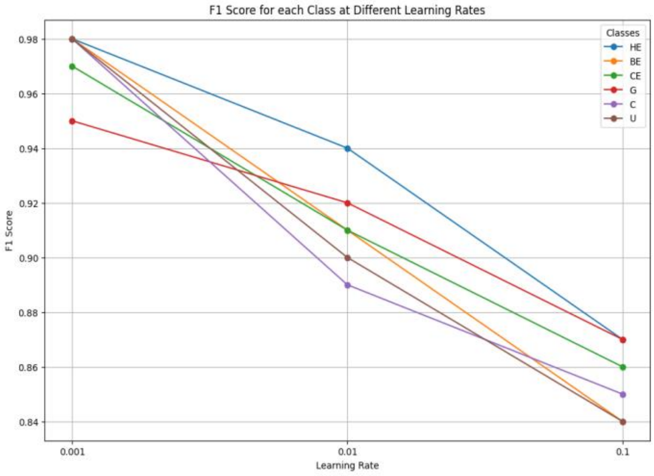

Fig. 12. F1 Score for each Class at Different Learning Rates

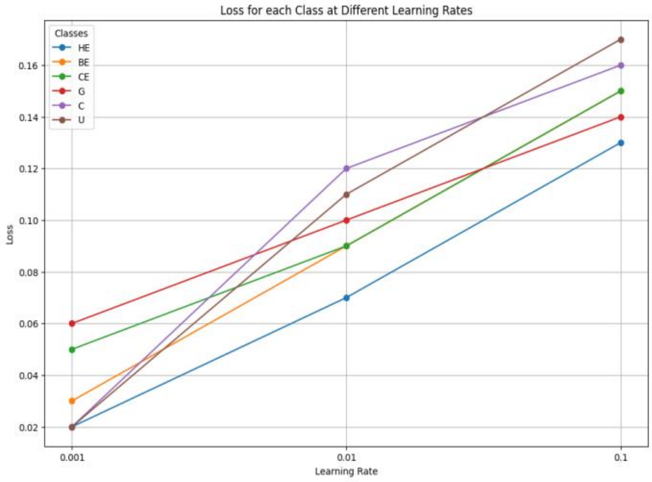

Fig. 13. Loss for each Class at Different Learning Rates

B. Comparison with State-of-the-Art Models

To put the hybrid model's results in perspective, they were compared to those of other, more established models. Figure 14 shows the accuracy comparison graph between the hybrid model and its competitors, and it demonstrates the hybrid model's impressive accuracy. HE and C are only two diseases for which the hybrid model has proven very effective at diagnosing, its practical relevance, and its promise for aiding medical practitioners. The graph depicts that the proposed model performs better than other models and has the highest accuracy, i.e. 98.13%.

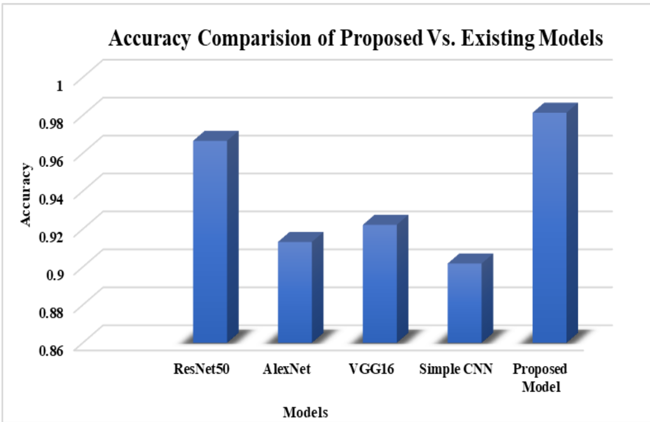

Fig. 14. Comparative analysis of accuracy achieved by different models and proposed model

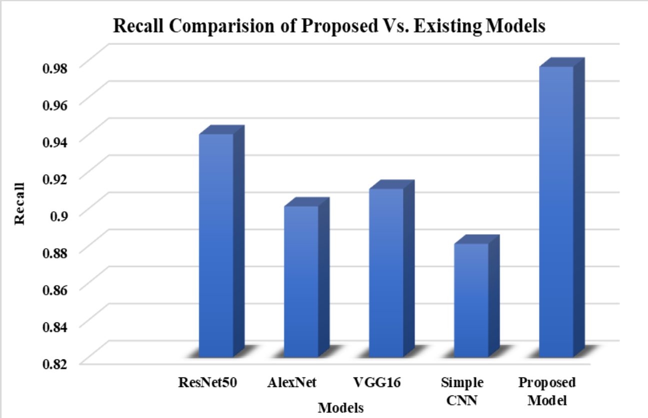

Fig. 15. Comparative analysis of recall achieved by different models and proposed model

Figure 15 shows a comparison graph of recall value for the proposed model and existing models, The result represents that the developed model performs better than other existing models and has the greatest recall value, i.e. 97.66%. Figure 16 shows the precision result of the comparison between the hybrid model and other existing DL models. As the graph in Figure shows, the hybrid model has the highest precision value, i.e., 98.09%, and performs better than other models.

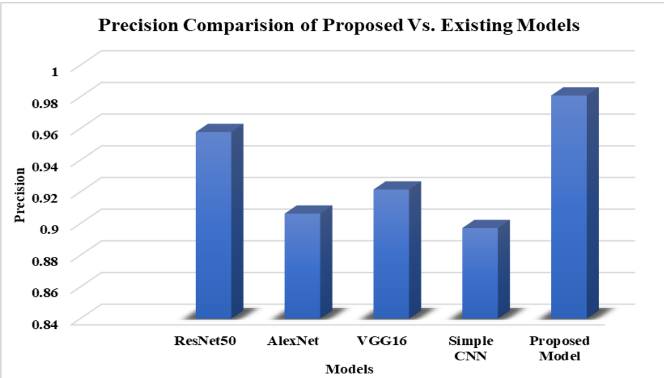

Fig. 16. Comparative analysis of precision achieved by different models and the proposed model

Figure 17 displays the bar chart that compares the F1 Measure of the proposed model against several existing models, ResNet50, AlexNet, VGG16, and Simple CNN. The graph shows that the proposed model has a better result than the other existing models with the highest F1 measure, which is 96.56%.

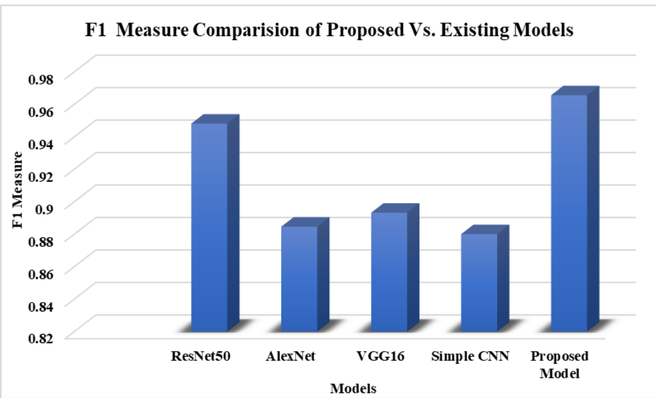

Fig. 17. Comparative analysis of F1 measure achieved by different models and proposed model

Overall, the proposed model demonstrates the highest F1 Measure compared to the existing models, indicating superior performance in terms of precision and recall. ResNet50 follows as the second-best performer, while AlexNet, VGG16, and Simple CNN show similar but lower performance levels.

Table 4 Comparative analysis of Proposed Model with Existing models

Model	Input Size	Accuracy	Recall	Precision	F1 Measure
Res-Net50	224X 224	96.65%	94.02	95.78	94.81
Alex Net	224 X 224	91.32%	90.13	90.63	88.46
VGG-16	224 X 224	92.22%	91.07	92.15	89.32
Simple CNN	224 X 224	90.18%	88.11	89.72	88.00
Prop. Model	224 X 224	98.13%	97.66	98.09	96.56

Table 4 illustrates the accuracy, recall, precision, and F1 measures obtained by different models such as ResNet50, Alex Net, AGG16, Simple CNN, and the newly proposed, hybrid model. The data, taken for prediction, defines the image size of 224x224.

The results show that the proposed, new hybrid model obtains the highest accuracy at 98.13%, highest recall at 97.66%, highest precision at 98.09%, and highest F1 Measure at 96.56% for detecting the illness or eye disease.

Figure 18 represents the comparison result of the proposed models with existing models for different performance measures. The graphical result shows that the proposed model has shown better results with the highest accuracy of 98.13%. The proposed model consistently outperforms the existing models (ResNet50, AlexNet, VGG16, and Simple CNN) in terms of accuracy, recall, precision, and F1 measure, indicating its superior performance and robustness. AlexNet shows the weakest performance, while ResNet50 performs well but still lags behind the proposed model. Overall, the hybrid model has shown promising results in all aspects.

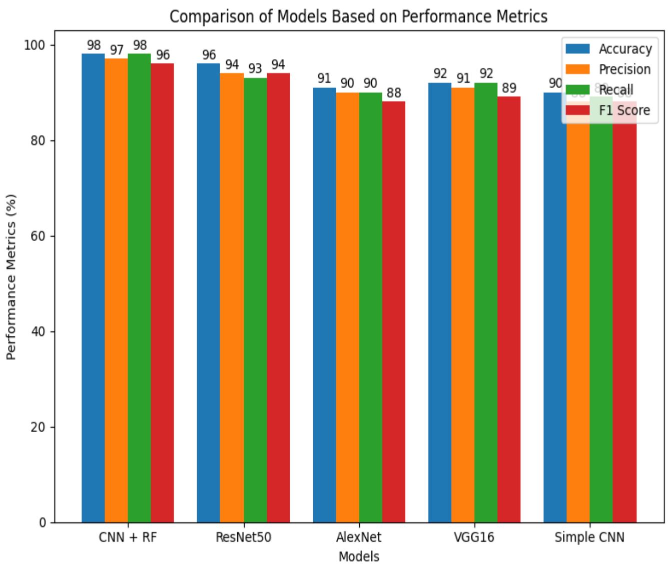

Fig 18. Comparison results of performance metrics for the proposed model vs. existing models

Table 5 Overall Comparison of the Proposed Model with Advanced Models

Model	Dataset	Accuracy (%)	Loss	Training Time	Model Complexity	Method
CNN + RF	Eye Disease Dataset	98%	0.05	2 hours	Medium	Hybrid CNN for feature extraction + RF
Deep CNN	Ocular Fundus Images	94%	0.15	3 hours	High	Deep CNN with transfer learning
VGG16	Eye Disease Dataset	93%	0.12	4 hours	High	Fine-tuning VGG16 on eye disease dataset
ResNet50	Retinal Images	95%	0.1	3.5 hours	High	Transfer learning with ResNet50.
Simple CNN	Eye Disease Dataset	88%	0.2	1 hour	Low	Basic CNN architecture with data augmentation
SVM with CNN	Ocular Disease Dataset	92%	0.18	2.5 hours	Medium	Extracted CNN features + SVM

Table 5 provides a structured comparison of your proposed CNN + Random Forest model with existing works, showing the strengths and advantages of the proposed model over previously established methods.

Table 6 Comparison results of the proposed model with existing work

Ref.	Accuracy	Recall	Precision	F1- Score	Models
[7]	88.3%	-	-	-	VGG16
[8]	80%	-	-	-	CNN
[10]	90%				CNN
[11]	92% 79%	-	-	-	ResNet &VGG16
[12]	84%	-	-	-	ResNet 152
[13]	87.5% 87.08%	-	-	91.3 85.42	SVM & CNN
[14]	89.3%	88.7	86.32	89.0	MLP
Props. Model	98.13%	97.66	98.09	96.56	CNN+RF

The author used the VGG16 [7] model and obtained an accuracy of 88.3% without any other classification. The author [8] used some basic classes of images for prediction that give 80% accuracy. As per the result presented in Table 5, the author used ResNet and VGG16 [11] models that show an accuracy of 92% and 79%. The model ResNet 152 [12] implemented by the author shows 84% accuracy. The author proposed the SVM and CNN [13] model for the prediction of multiclass eye disease that obtained an accuracy of 87.5% and 87.08%. Similarly, the author achieved 89.3% accuracy using DenseNet121 and Inception-ResNetV2 [14], and the features are extracted using the MLP model. As presented in Table 6, the proposed CNN+RF model outperforms with 98.13% accuracy as the previous work done by other authors. Furthermore, the result also shows that the proposed model is suitable for classifying and predicting these kinds of diseases and analysis. The proposed Model (Using CNN+RF), achieved the highest accuracy of 98.13%, recall of 97.66%, precision of 98.09%, and F1-Score of 96.56%.

Table 7 provides an in-depth comparison of how different learning rates affect the performance of the proposed model across various eye disease classes. Learning rate is a crucial hyperparameter that influences the speed and quality of model training. The model achieves its highest accuracy at a learning rate of 0.001, consistently exceeding 96% across all disease categories. Lower learning rates allow for finer adjustments during training, reducing the chances of overshooting optimal weights. Precision remains above 97% across all classes at 0.001, ensuring minimal false positives. Recall values indicate that the model effectively identifies true cases, with Uveitis (UT) and Hypertensive Retinopathy (HE) achieving nearly 99% accuracy.

Table 7 Classification Performance Comparison of different model with Proposed Model

Model	Classification Speed	Training Time	Model Complexity	Robustness
ResNet50	Fast	Medium	High	High
AlexNet	Moderate	Medium	Medium	Medium
VGG-16	Slow	High	High	Medium
Simple CNN	Fast	Low	Low	Low
Proposed Model (CNN+RF)	**Very Fast**	**Medium**	**Medium**	**High**

Since F1-Score is the harmonic mean of Precision and Recall, it demonstrates model stability. The model maintains a near-perfect F1-Score across all classes, with slight variations depending on class complexity. The lowest loss values (below 0.03) are observed for HE, C, and U, signifying that the model is well-optimized for these disease categories. Loss values slightly increase for Glaucoma (G) and Crossed Eyes (CE), suggesting potential areas for further fine-tuning. If the learning rate is too high, the model might converge quickly but risk overfitting or unstable performance. If the learning rate is too low, training might take significantly longer while providing only marginal improvements.

The proposed model, which integrates CNN and Random Forest (RF), outperforms the other models across all evaluated metrics, demonstrating superior accuracy, recall, precision, and F1-Score. This indicates that the combination of CNN and RF is particularly effective for the given classification task, making it the most robust model among those compared. The high-performance metrics suggest that this model can accurately identify true positives and negatives, making it highly reliable for practical applications.

C. Discussion

The hybrid model works on the combined properties of the CNN and RF Models. The CNN module effectively pulls complex characteristics from the ocular images, encompassing spatial patterns, textures, and structures, that are essential characteristics for precise illness detection. The features are subsequently directed toward the RF component, wherein a collective of decision trees work together to generate accurate and comprehensible categorization results. The hybrid model is strong for a wider variety of ocular diseases, as seen by the acquired overall accuracy of 98.13%. Metrics such as precision, recall, and F measures for each illness category highlight the model's accuracy in case identification and its capacity to reduce false positives and negatives. Notably, the model's improved accuracy in C classification indicates its prowess in correctly identifying this ailment. Understanding the model's specific difficulties requires looking at the confusion matrix. Inherent similarities in ocular features make distinguishing between BE and G cases difficult, as seen by the confusion between the two. This information is crucial for shaping future adjustments and making the model more sensitive to the nuances at work.

V. Conclusion and Future Scope

The proposed study used a hybrid DL model to detect and classify eye disorders, and the results were presented in Culmination. The study relied on a sizable collection of eye images gathered from the Kaggle repository and augmented with additional data. Integrating CNN and RF, the hybrid model accurately recognized and categorized six ocular pathologies: HE, BE, C, CE, G, and UT. The empirical data shows that the hybrid model has an overall accuracy of 98.13 percent. The model's performance for different diseases was thoroughly analyzed using different metrics such as precision, recall, and F1-score. The confusion matrix visualization demonstrated the model's ability to discriminate between various eye states, highlighting its strengths and weaknesses. In addition, in the context of G and UT categorization, the comparison with state-of-the-art models demonstrated the hybrid model's potential as a competitive option. The results of this study help push ophthalmic medical image analysis further. If the hybrid model is effective, it could dramatically improve the diagnosis and classification of eye diseases, leading to more targeted therapies. The implications extend beyond the dominion of technology, offering doctors a comprehensive array of resources to enhance patient care. As this investigation ends, the hybrid model is at the crossroads of technological advancement and practical use in healthcare. It represents a game-changing force in the management of ocular health because it combines the feature extraction powers of DL with the decision-making proficiency of RF. The hybrid model has the potential to become a crucial tool for improving ocular health through further development and incorporation into medical practices.

Data Availability

Raw data supporting the conclusions of this investigation are accessible from the corresponding author upon reasonable request.

References

[1]. T. Nazir, A. Irtaza, A. Javed, H. Malik, and R. A. N. Dildar Hussain, “Retinal image analysis for diabetes-based eye disease detection using deep learning,” Appl. Sci., 10(18), 6185–6172, 2020.

[2]. Abbas, Qaisar, Imran Qureshi, Junhua Yan, and Kashif Shaheed. "Machine learning methods for diagnosis of eye-related diseases: a systematic review study based on ophthalmic imaging modalities." Archives of Computational Methods in Engineering 29(6) 6 (2022): 3861-3918.

[3]. K. Lin, R. Chen, J. Chen, P. Lu and F. Yang, “Multi-View Block Matrix-Based Graph Convolutional Network”, Engineering Letters, 32(6), 1073-1082, 2024.

[4]. Jain, L., Murthy, H. S., Patel, C., & Bansal, D., Retinal eye disease detection using deep learning. In 2018 Fourteenth International Conference on Information Processing (ICINPRO), 1-6. IEEE, 2018.

[5]. T. Naresh et al., “Early detection and classification of tomato leaf disease using high-performance deep neural network,” Sensors, 21(23), 1050–1061, 2021.

[6]. S. Shekhar, H. Garg, R. Agrawal, S. Shivani, and B. Sharma, “Hatred and trolling detection transliteration framework using hierarchical LSTM in code-mixed social media text,” Complex Intell. Syst., 5(11), 1–14, 2021.

[7]. Sarki, R., Ahmed, K., Wang, H., & Zhang, Y. (2020). Automated detection of mild and multi-class diabetic eye diseases using deep learning. Health Information Science and Systems, 8(1), 32.

[8]. K. Prasad, P. S. Sajith, M. Neema, L. Madhu, and P. N. Priya, “Multiple eye disease detection using Deep Neural Network,” in IEEE Region 10 Conference, 2148–2153, 2019.

[9]. Nazir, Tahira, Marriam Nawaz, Junaid Rashid, Rabbia Mahum, Momina Masood, Awais Mehmood, Farooq Ali, Jungeun Kim, Hyuk-Yoon Kwon, and Amir Hussain., “Detection of diabetic eye disease from retinal images using a deep learning based CenterNet model,” Sensors, 21(16), 5283–5290, 2021.

[10]. P. Chakraborty and C. Tharini, “Pneumonia and eye disease detection using convolutional neural networks,” Eng. Technol. Appl. Sci. Res., 10(3) 5769–5774, 2020.

[11]. Amin, R., Ahmed, A., Hasan, S.S.U. and Akbar, H., 2022. Multiple Eye Disease Detection Using Deep Learning. Foundation University Journal of Engineering and Applied Sciences (HEC Recognized Y Category, ISSN 2706-7351), 3(2), pp.14-26.

[12]. N. Badah, A. Algefes, A. AlArjani, and R. Mokni, “Automatic eye disease detection using machine learning and deep learning models,” in Pervasive Computing and Social Networking: Proceedings of ICPCSN, 773–787, 2022.

[13]. Pahuja, Rahul, Udit Sisodia, Abhishek Tiwari, Siddharth Sharma, and Preeti Nagrath. "A Dynamic approach of eye disease classification using deep learning and machine learning model." In Proceedings of Data Analytics and Management: ICDAM 2021, Volume 1, pp. 719-736. Springer Singapore, 2022.

[14]. R. H. Paradisa, A. Bustamam, W. Mangunwardoyo, A. A. Victor, A. R. Yudantha, and P. Anki, “Deep Feature Vectors Concatenation for Eye Disease Detection Using Fundus Image.,” Electronics, 11(23),1452–1463, 2022.

[15]. D. Shamia, S. Prince, and D. Bini, “An Online Platform for Early Eye Disease Detection using Deep Convolutional

Neural Networks," in International Conference on Devices, Circuits and Systems, 388–392, 2022.

[16]. Li, Feng, Hua Chen, Zheng Liu, Xuedian Zhang, and Zhizheng Wu. "Fully automated detection of retinal disorders by image-based deep learning." Graefe's Archive for Clinical and Experimental Ophthalmology 257 (2019): 495-505.

[17]. R. Sharma and V. Kukreja, "CPD: Faster RCNN-based DragonBall Comic Panel Detection," in International Conference on Communication Systems and Network Technologies, 2023, pp. 786–790.

[18]. P. Singla, Niharika, R. Jain, R. Sharma, V. Kukreja, and A. Bansal, "Deep Learning Based Multi-Classification Model for Rice Disease Detection," in International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), 2022, pp. 1–5.

[19]. Chelaramani, Sahil, Manish Gupta, Vipul Agarwal, Prashant Gupta, and Ranya Habash. "Multi-task learning for fine-grained eye disease prediction." In Asian Conference on Pattern Recognition, pp. 734-749. Cham: Springer International Publishing, 2019.

[20]. Singh, S., & Mittal, S., "Pandemic Outbreak Prediction using Optimization-based Machine Learning Model," in 2023 3rd

[21]. International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), 2023, pp. 154-159. IEEE.

[22]. Singh, S., Mittal, S. and Singh, S., "Analysis and Forecasting of COVID-19 Pandemic Using ARIMA Model". In 2023 3rd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), 2023, pp. 143-148. IEEE.

[23]. Tan, N. M., Liu, J., Wong, D. W., Zhang, Z., Lu, S., Lim, J. H., ... & Wong, T. Y. Classification of left and right eye retinal images. In Medical Imaging 2010: Computer-Aided Diagnosis, 7624(1), 913-920). SPIE.

[24]. Isha Baokar, and Lili He, "Memristive Neural Networks Application in Predicting of Health Disorders," Lecture Notes in Engineering and Computer Science: Proceedings of The International Multi Conference of Engineers and Computer Scientists 2023, 5-7 July, 2023, Hong Kong, pp94-99