Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Hybrid Deep Learning Approach for Eye Disease Detection with Integrating Adaptive CNN Layers for Improved Performance of Image Classification

Piyush Jain[1], Dr. Deepak Motwani[2] and Dr. Pankaj Sharma[3]
[1]*PhD Scholar, Amity University, Gwalior, Madhya Pradesh, India, piyushjain.1987@gmail.com*
[2]*Associate Professor, CSE Amity University Gwalior, India, dmotwani@gwa.amity.edu*
[3]*Computer Science and Engineering, Eshan College of Engineering, Mathura, India, topankajsharma126@gmail.com*

ARTICLE INFO

Received: 24 Oct 2024

Revised: 25 Nov 2024

Accepted: 18 Dec 2024

ABSTRACT

Essential for the proper functioning of eyesight, the eye is a crucial organ. If the retina sustains permanent damage from an eye disease, the patient may have severe blurred vision or perhaps become blind. The use of AI in eye disease categorization (EDC) helps doctors better serves their patients. This study's overarching goal is to construct an EDC model by means of deep learning (DL). This research details a process for detecting eye diseases using DL models that use binary and multi-class classification methods. Leveraging the APTOS 2019 dataset, the Adaptive Hybrid Net-CNN model integrates InceptionV3, MobileNetV3Small, ResNet101, and adaptive CNN layers to enhance feature extraction and classification performance. The preprocessing pipeline involves image auto-cropping, resizing, RGB conversion, and sharpening to ensure high-quality inputs. For binary classification, labels are binary-encoded, achieving accuracies of 99.94% during training and 98.45% during testing, with strong performance metrics, including an F1-score of 0.98. For multi-class classification, labels are consolidated into three categories, and SMOTE is used to balance the dataset also K-fold cross-validation further validated the models. An F1-score of 0.97 and an accuracy of 99.28% on training data and 96.36% on testing data were achieved by the model, which was trained utilising categorical cross-entropy loss. Evaluation is conducted through classification reports, ROC and precision-recall curves, and accuracy/loss graphs, offering a thorough understanding of model performance. The comparative analysis shows that the proposed model works well in comparison to existing (GoogleNet_SVM, Hybrid_SVM, and ResNet-18_SVM) models in terms of all performance measures. The proposed model demonstrates effective and scalable eye disease detection, delivering both binary and multi-class classification solutions. The outcomes show how important the suggested EDC model is for disease classification in complicated fundus images of the eye.

Keywords: Diabetes Retinopathy, Eye Disease, Medical Imaging, Deep learning, CNN, Adaptive Hybrid Net, APTOS 2019.

INTRODUCTION

Healthcare systems worldwide recognize the critical need for early identification and diagnosis of diabetes-related disorders for better treatment and care. In many people, diabetes is one of several co-occurring disorders[1]. A metabolic disorder causes blood sugar levels to remain consistently high and eventually impacts several bodily functions[2]. Nearly 463 million individuals throughout the globe were living with diabetes in 2019[3]. A diabetes pandemic has recently emerged in developing nations like India and China[4]. The International Diabetes Federation predicts that this trend will continue from 2030 to 2045, with the worst impacts seen in India and China[5]. In 2019, the 3 nations with a highest rate of diabetes were the United States, China, and India. Further, it seems that males have a higher prevalence of diabetes (9.0%) compared to women (7.9%). There is no denying that the eyes are among the many bodily organs severely affected by diabetes[6]. A condition that may cause blindness in the eyes, DR, is more common in those with diabetes. DR is a disease of the eye that may cause permanent damage to the retina and, in extreme cases, blindness[7]. The eye is a vital organ for everyday life. Serious eye illnesses have the potential to

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

permanently harm the retina[8]. The eye is a vital organ for everyday life. Serious eye illnesses have the potential to permanently harm the retina[9].

There was a remarkable recovery rate when eye disorders were detected early. It takes a lot of time for traditional diagnostic procedures that mostly depend on ophthalmology specialists to complete the process. There is a wealth of medical images available for use in creating more precise diagnostic tools, made possible by recent technical developments in imaging methods. The analysis of medical images is being greatly assisted by deep learning (DL) models[10, 11]. DL algorithms can automatically learn the signs of a variety of eye diseases, including DR, AMD, DME, and others, using ocular photo datasets. AI makes heavy use of deep learning (DL), a kind of ML[12]. DL generally performs better than all conventional image analysis methods and can reliably extract features by input data for segmentation or classification. While DL approaches demand a large amount of data for training, it do not require the hand-crafted features to be extracted[13-15]. An example of a DL approach, CNN, is a popular, powerful, and effective tool for picture analysis[16, 17]. CNNs, which are excellent at automatically extracting features and learning, are among its most successful methods. CNN greatly reduces processing loads by using convolution kernels to analyze images in narrow perceptual domains. Convolution layers are layered in modern DNNs to further deepen CNNs, as shown in designs like VGGnet[18], ResNet[19], and GoogLeNet [20, 21]. CNNs have shown great promise in a number of domains, most notably image classification and voice recognition[20]. This paper introduces a hybrid DL approach for eye disease detection using the Adaptive Hybrid Net-CNN model, combining InceptionV3, MobileNetV3Small, ResNet101, and adaptive CNN layers. The methodology involves preprocessing images through auto-cropping, resizing, and sharpening, followed by binary and multi-class classification tasks. The APTOS 2019 dataset is used for binary classification, while class imbalances in multi-class detection are addressed with techniques like SMOTE. Adam is utilized to optimize a model, and measures like F1score, recall, precision, and ROC curves are utilized to assess performance. K-fold cross-validation is used to assure robustness.

A. Research Motivation and Significance

The rising incidence of vision-related diseases, like DR, glaucoma, and macular degeneration, the three main causes of blindness worldwide, is driving research into the development of methods for detecting eye diseases using fundus images. Traditional manual procedures are laborious and error-prone, but early identification is crucial for avoiding visual loss. Leveraging deep learning models to analyze fundus images can improve diagnostic accuracy, speed up the process, and facilitate early interventions. In the long run, this strategy might improve patient outcomes and lessen the impact of blindness by making ophthalmic treatment more accessible, affordable, and efficient.

B. Novelty and Contribution

The novelty of this work lies in the development of the Adaptive Hybrid Net-CNN model, which combines multiple state-of-the-art architectures—InceptionV3, MobileNetV3Small, and ResNet101—along with adaptive CNN layers for enhanced feature extraction and classification. The use of a comprehensive preprocessing pipeline, including SMOTE for handling class imbalances in multi-class detection, further distinguishes this approach. Additionally, the model's robust evaluation metrics and application of K-fold cross-validation provide a reliable framework for real-world deployment in eye disease detection. This hybrid approach offers an innovative solution for accurate and scalable diagnostics. The following proposed research contributions of this paper are:

- The integration of multiple advanced architectures (InceptionV3, MobileNetV3Small, ResNet101) with adaptive CNN layers enables the model to capture diverse features from retinal images, significantly improving classification accuracy and robustness.
- The development of a robust preprocessing pipeline, including auto-cropping, resizing, RGB conversion, and sharpening filters, ensures high-quality input data, addressing common image inconsistencies and enhancing model performance.
- The research addresses class imbalance in multi-class classification through techniques like SMOTE, which improves model generalization and ensures accurate results even in the presence of unequal class distributions.
- The technique incorporates a full suite of assessment measures to evaluate a model's performance from several perspectives. These metrics include accuracy, F1-score, precision, recall, and sensitivity. Additionally, ROC and precision-recall curves are used.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- The application of K-fold cross-validation for multi-class detection ensures the robustness of the model, preventing overfitting and guaranteeing reliable results across various subsets of the dataset.
- The suggested model is built to be flexible and scalable, making it ideal for use in clinical settings. It provides a reliable and applicable method for identifying eye illnesses, which advances the field of automated disease diagnosis and medical picture analysis.

C. Structured of the paper

An overview of the paper is as follows: A brief summary of pertinent papers on deep learning-based DR detection is given in Section II, while Section III details the materials and techniques that have been suggested. Section IV provides a comprehensive account of the results and experiments. Section V provides an overview of the findings and makes recommendations for further research.

LITERATURE REVIEW

In cases of DR, automated systems can help discover cases early and avoid irreversible blindness by eliminating the need for specialised expertise and patient fatigue associated with human grading. Researchers have investigated many methods for the categorization of DR in an effort to resolve these issues. This section provides an overview of the models utilised in this field of study.

Ruminaviciute and Bernataviciene (2023). The goal of this study is to provide a dependable approach for automatically assessing picture quality by assembling a set of techniques to analyze both the images' deep features and other indications. The ensemble demonstrated the viability of ML approaches for trustworthy automated image analysis by attaining high accuracy rates of 0.9146 and 0.9845 on two datasets of digital fundus images[21].

Abini and Priya (2023) The main goal of this study is to use the ResNet-50-CNN architecture to categorize fundus pictures into various severity levels. The APTOS 2019 dataset from Kaggle may be used to test the algorithm's efficacy. Training two-class models on the APTOS dataset yields an accuracy of 96% and verifying them yields an accuracy of 95%. On the enlarged APTOS dataset, the training accuracy for multi-class classification is 77% and the validation accuracy is SO%[22].

Chavan and Pete (2023) Retinal fundus images combined with transfer learning models allow for the early diagnosis of several illnesses. The InceptionV3 transfer learning model and the strong VGG 16 have been very beneficial to medical picture analysis. The VGG 16 model has an AUROC curve of 0.913, and the Inception V3 model has an AUROC of 0.942. Compared to VGG 16, Inception V3 performs better on the RFMiD dataset. RFMiD is composed of color fundus images of the retina[23].

Sharma et al. (2024) this research compared utilizing ML and deep/transfer learning to categorize fundus images. Various model architectures are used, including SVM, Gradient Boosting, Efficient Net, Efficient-Net B4 and B7, MLP, CNN, ResNet, VGG-19, and DNN, among others. Their strategies outperform a single model with results ranging from 84% to 96% accuracy, respectively[24].

Oulhadj et al. (2024) paper introduces a new automatic technique for identifying the degree of severity of DR, based on a new hybrid DL approach (DenseNet121, Xception, and EfficientNetB3) with a pre-processing step. They demonstrated the efficacy of their method by doing tests on the APTOS dataset, where it achieved an outstanding 86% accuracy rate[25].

Uttakit et al. (2024) the paper studied the performance of a DL-pretrained CNN called EfficientNetB7 has been utilized. The findings showed that the model had a sensitivity level of 88.17%, a specificity level of 94.08%, and an accuracy level of 88.17% when predicting the three classes [26].

Malik et al. (2024) propose a VGG-19-based classifier for automated DR detection and classification. Through transfer learning and fine-tuning, they adapt the VGG-19 model for DR classification, utilizing its learned representations to enhance their model's performance. The experimental findings show a promising accuracy of 64.5% for four-class classification and the effectiveness of VGG-19-based classifier in identifying four stages of DR[27].

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Table 1 summarizes the methodologies, datasets, results, limitations, and future work for each of the papers for fundus image classifications using various deep-learning techniques.

Table 1. Previous work on fundus photograph classification with various techniques

References	Methodology	Dataset	Results and Key Analysis	Limitations	Future Work
Ruminaviciute et al. (2023)	Ensemble of methods for automated image quality assessment using deep features and indicators.	Two digital fundus image datasets	Accuracy: 0.9146 and 0.9845. Demonstrates potential for ML in telemedicine for image quality assessment and ensuring accurate diagnoses.	Limited to two datasets and may not generalize well to others.	Explore a larger variety of datasets and expand image quality indicators.
Abini and Priya (2023)	ResNet-50-CNN architecture for classifying fundus images into severity levels.	Kaggle APTOS 2019	Accuracy for binary classification: 96% (train), 95% (validate); for multiclass classification: 50% (train), 77% (validate).	Lower performance for multiclass classification.	Improve multiclass classification accuracy and optimize the model.
Chavan and Pete (2023)	Transfer learning models (VGG16, InceptionV3) for multi-disease detection using retinal fundus images.	RFMiD retinal fundus image dataset	AUROC: VGG16: 0.913, Inception V3: 0.942. Inception V3 outperforms VGG16.	Limited to two models and may not capture all disease types.	Incorporate additional models and evaluate performance on diverse datasets.
Sharma et al. (2024)	Comparison of various ML and DL models (SVM, Gradient Boosting, CNN, ResNet, etc.) for DR classification.	Fundus image data	Accuracy: 84%-96% across models. ML-DL-TL hybrid methods outperform single models.	Limited exploration of new model architectures.	Test with additional datasets and explore other DL architectures.
Oulhadj et al. (2024)	Hybrid deep learning approach (DenseNet121, Xception, EfficientNetB3) for DR severity detection.	Kaggle APTOS	Accuracy: 86%. Demonstrates hybrid DL model with pre-processing for enhanced image quality.	Only validated on one dataset (APTOS).	Expand to multiple datasets and improve hybrid model performance.
Uttakit et al. (2024)	EfficientNetB7 for glaucoma severity detection using transfer learning.	Fundus photographs	Accuracy: 88.17%, Specificity: 94.08%, Sensitivity: 88.17%. Comparable to state-of-the-art methods	Focus on limited glaucoma severity categories.	Explore additional glaucoma stages and test with more diverse datasets.
Malik et al. (2024)	VGG-19-based classifier with transfer learning for DR detection and classification.	Retinal fundus images	Accuracy: 64.5% for four-class classification. Demonstrates effectiveness of VGG-19 in identifying DR stages.	Limited to four-class classification and may not generalize to other conditions.	Explore additional disease stages and refine VGG-19 model for DR.

METHODOLOGY

The aim of this work is to design and implement a hybrid deep learning approach for eye disease detection by integrating adaptive CNN layers for improved performance of image classification. The proposed methodology for

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

eye disease detection, encompassing both binary and multi-class classification, leverages the Adaptive Hybrid Net-CNN model, integrating InceptionV3, MobileNetV3Small, ResNet101, and adaptive CNN layers for robust feature extraction and classification. The preprocessing pipeline ensures high-quality inputs by auto-cropping, resizing to 224x224, converting images to RGB, and applying sharpening filters. For binary classification, the APTOS 2019 dataset is utilized, with labels binary-encoded and split into an 80:20 train-test ratio, while for multi-class detection, labels are consolidated into three categories to simplify classification. Class imbalances in the multi-class approach are addressed using SMOTE or similar techniques. Training for 50 epochs with 132 batches, both models utilize the Adam optimizer with a learning rate of 0.001 and, depending on the situation, apply binary or categorical cross-entropy loss functions. Performance evaluation includes metrics like F1score, precision, recall, sensitivity, and tools such as ROC and precision-recall curves, with K-Fold cross-validation applied for multi-class detection to ensure robustness. This integrated methodology emphasizes preprocessing, hybrid architecture design, and comprehensive evaluation to deliver accurate and generalizable results for eye disease detection. Figure 1 displayed an entire procedure of the proposed approach for classifying fundus photos for deep learning-based eye disease diagnosis.

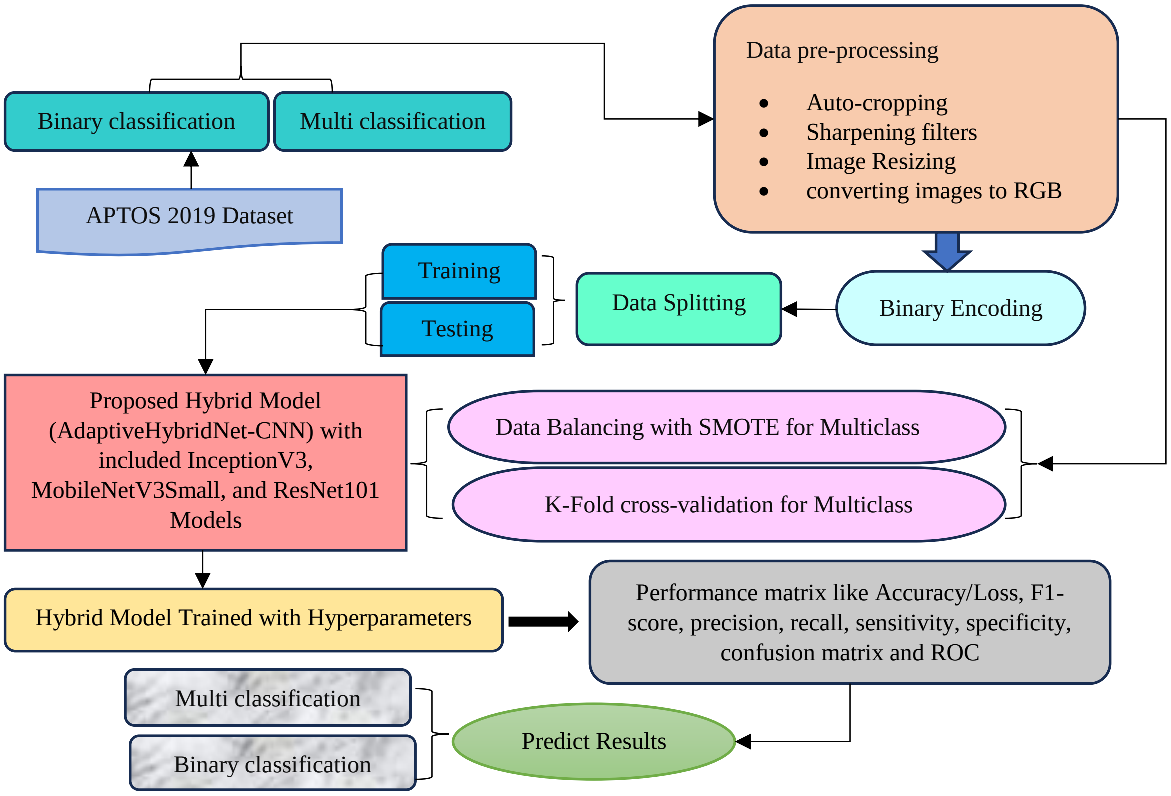


Figure 1. Proposed Flowchart for fundus photographs classification for eye disease detection based on deep learning

The methodology steps are discussed below and clearly described the flowchart in Figure 1:

A. Data Collection

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

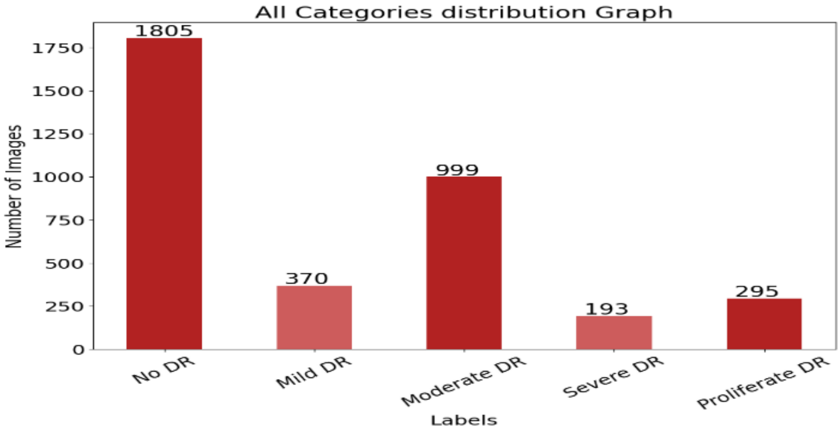


Figure 2. Bar Graph of Multi-Class Distribution

In this paper, the dataset utilized for both binary and multi-class eye disease detection is the APTOS 2019 dataset in the Kaggle challenge, which contains fundus images of the retina along with corresponding labels. Figure 2 shows that 3662 retinal samples make up this database. These samples are sorted into five groups: negative, mild, moderate, proliferative, and severeDR. There are a total of 1805 samples assigned to the negative DR class, 370 to the moderate DR class, 999 to the severe DR class, 193 to the proliferative DR class, and 295 images to a last class.

B. Data Preprocessing

Figure 3 shows the images taken of the retina before and after preprocessing, which highlights the importance of image quality control in enhancing the network's performance and ensuring that all images are consistent while also enhancing their features[28]. The following data preprocessing is discussed below and applied to APTOS 2019 dataset:

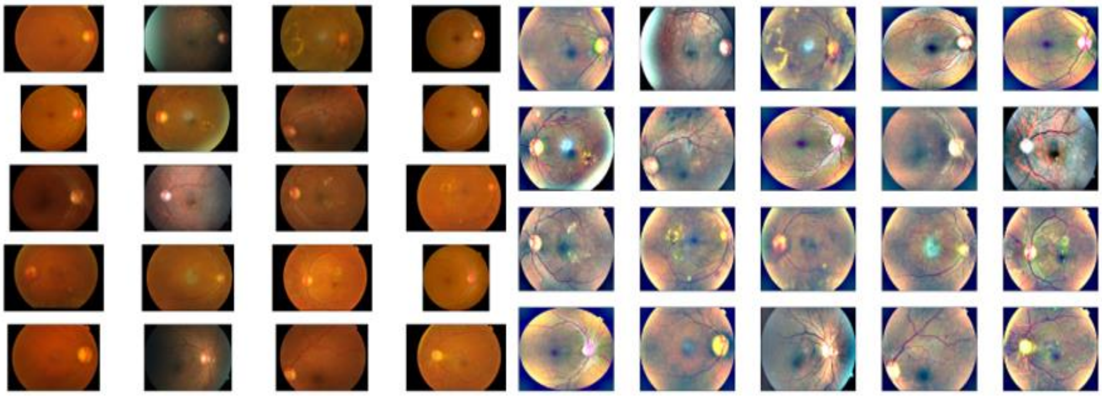


Figure 3. After and Before Data Preprocessing Images

- **Color Conversion:** Convert images by BGR (OpenCV default) to RGB color space to standardize for further processing and model input.
- **Auto-Cropping (if needed):** Apply an auto-cropping technique to remove any unnecessary borders around the images to focus on the relevant area (e.g., the retina).
- **Resizing:** Ensuring consistency and meeting the input criteria of the model, all images are shrunk to a constant size of 224x224 pixels.
- **Sharpening:** A sharpening filter is applied to enhance image details and improve the quality of feature extraction. This is done using the cv2.addWeighted method for better clarity of important features.
- **Label Mapping:** For binary classification, labels are directly mapped as 0 or 1 based on the presence or absence of disease. For multi-class classification, labels are consolidated into three categories:
  - Class 0: For diagnosis 0.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- Class 1: For diagnoses 1 and 2.
- Class 2: For diagnoses 3 and 4.

C. Deep Learning Classification Models

The use of DL has completely transformed medical image analysis. This research provides an Adaptive Hybrid Net-CNN model-based hybrid DL strategy for eye disease identification. The model combines InceptionV3, MobileNetV3Small, ResNet101, and adaptive CNN layers for enhanced feature extraction and classification tasks on APTOS 2019dataset is used for binary classification, while class imbalances in multi-class detection are addressed with techniques like SMOTE. Model performance is evaluated when Adam makes optimizations using metrics like F1score, recall, precision, and ROC curves. K-fold cross-validation is utilized to guarantee robust findings. The following pre-trained models are:

1) InceptionV3

The InceptionV3 is designed to be both accurate and efficient in terms of processing resources. When it comes to CNN architectures, InceptionV3 is a huge leap forward, particularly for picture identification and classification. Google researchers created InceptionV3[29] and it is famous for its revolutionary usage of inception modules, which let the network capture and process data at several spatial scales. This model has a complex structure that includes various kernel sizes, permitting a model to extract coarse-grained features by input images. Furthermore, the structure employs auxiliary classifiers at some stage in training to mitigate the vanishing gradient problem. Its versatility, performance, and superior overall performance have made InceptionV3 of significant importance within the field of gaining deep knowledge of computer vision. The operation of an inception module in inceptionV3 involves concatenating feature maps obtained from different convolutional filters as shown in Equation (1).

$$Y^{(l)} = concat(Conv1 \times 1, Conv3 \times 3, Conv5 \times 5, MaxPool \tag{1}$$

Here, $Y(l)$ represents an output feature map and Concat denotes the concatenation operation.

2) MobileNetV3Small

The third version of the MobileNet series, MobileNet-v3, is an architecture for CNNs that has considerably improved accuracy and efficiency compared to its predecessors. Achieving a compromise between accuracy and efficiency, MobileNet-v3 employs lightweight construction blocks and incorporates sophisticated features like squeeze-and-excitation modules. The model's portability was prioritized in order to meet the demands of embedded vision applications and mobile devices used in the automobile industry[30]. MobileNetV3Small, designed for efficiency on mobile and resource-constrained devices, contributes lightweight and computationally efficient feature extraction capabilities, making the model scalable and practical.

3) ResNet101

The strategy's importance was demonstrated by means of the ImageNet-pretrained ResNet101 residual network. Eye illness categorization is one of several transfer learning projects that have made use of this architecture, which is also employed as a backbone network[31]. Standard CNNs with stacked residual blocks are called ResNets. Every block can be expressed Equation (2) and (3) in the following way[32]:

$$y_I = x_I + \mathcal{F}(x_I; W_I) \tag{2}$$

$$x_{I+1} = f(y_I) \tag{3}$$

where $x_I$and x_{I+1} Representing the n-th residual block's input and output, F stands for a residual function with a corresponding weight, Wl and stands for an activation function of a ReLu. Activation function, batch normalization, and convolutional layers are all part of function F. For ResNet101, there are a total of 33 residual blocks.

4) Adaptive CNN layers

This paper highlights the use of adaptive CNN layers in a hybrid deep-learning approach for eye disease detection. These adaptive layers are designed to dynamically optimize feature extraction by adjusting to varying complexities in medical images, ensuring a more precise and robust representation of critical patterns. By integrating adaptive

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

CNN layers into the broader architecture, the model enhances its ability to capture subtle variations in input data, significantly improving classification performance. This adaptability makes the model suitable for both binary and multi-class classification tasks, supporting advanced medical diagnostics with greater accuracy and generalization.

D. Model Training (Adaptive Hybrid Net-CNN model)

The Adaptive Hybrid Net-CNN model is trained using the Adam optimizer, which is famous for its efficacy in training DL models. Throughout training, the learning rate will be kept at 0.001 to guarantee steady and progressive convergence. The loss function that is most often used for binary classification is binary cross-entropy. However, when dealing with multi-class problems, categorical cross-entropy is preferable since it may punish wrong predictions using the softmax output. Training the model for 50 epochs with a batch size of 132 ensures that the model undergoes sufficient iterations to acquire complicated feature information from the images.

These multiclass pre-trained models were fitted using categorical cross-entropy loss functions. Perform the computation of category cross entropy with N labels[33] Equation (4):

$$L(\hat{y}, y) = -\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n} y_{ij} * \log(\hat{y}_{ij}) \tag{4}$$

One major advantage of categorical cross entropy is that it is very easy to calculate the gradient of $\boldsymbol{L(\hat{y}, y)}$.

Since this is a binary classification task (disease vs. no disease), the binary cross entropy loss is appropriate. A discrepancy among the expected probability and the actual labels is measured quantitatively. The goal of binary classification issues is to divide data into two groups, such as benign and malignant in breast cancer diagnosis. One popular loss function for such problems is binary cross entropy. It finds the discrepancy among a model's projected probabilities (y ̂) and an actual labels (y). A function penalizes incorrect predictions, especially when the confidence is high but incorrect. For a single instance as Equation (5):

$$L = -\frac{1}{N}\sum_{i=1}^{N}(y log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{\bar{y}}_i)) \tag{5}$$

Where:

- y : true label (0 or 1).
- $\hat{y}_i$: Predicted probability of the positive class (0 ≤ $\hat{y}_i$).

E. Model Evolution

The Adaptive Hybrid Net-CNN model is evaluated employing accuracy, F1-score, specificity, recall, sensitivity, and precision, alongside tools like confusion matrix, ROC, and precision-recall curves to assess performance and robustness for both binary and multi-class classification tasks.

RESULTS ANALYSIS AND DISCUSSION

The results of the proposed AdaptiveHybrid Net-CNN model's experiments on the APTOS 2019 dataset for fundas image classification are displayed below. The model is implemented using a Python simulation tool that is based on binary and multiclass classification. The following sections provide the experimental simulation, evaluation Metrix like precision, accuracy, f1-score, recall, specificity, ROC and PR, results of proposed model and at last discuss the comparative analysis that shows the proposed model works well in compared to other existing models on APTOS 2019 dataset.

A. Experimental Setup

The experimental setup for the fundus image classification system was optimized for the efficient handling of resource-demanding tasks. A local workstation boasting an Intel i7 CPU, 16 GB RAM, and an NVIDIA GTX 1660 Ti GPU was utilized to create the system, with cloud-based processing handled by Google Colab. Python was the primary programming language, with the project developed using Jupyter Notebook and Colab environments. Data preparation, exploratory data analysis, and creating DL models for the categorization of fundus images were accomplished using essential libraries like Pandas, NumPy, Matplotlib, Seaborn, scikit-learn, Keras, and TensorFlow.

B. Evaluation Metrics

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

ROC, F1Score, specificity, recall, specificity, acc-uracy, precision, precision-recall, and other metrics are used to assess the proposed approach. To find out whether the planned study would be beneficial, this research uses many performance assessment criteria. A confusion matrix is created during the procedure, which may be used to score the method's efficacy based on a number of various metrics. The confusion matrix produced during the identification task's testing is the foundation of the majority of methods. Specificity, like predictability, is the proportion of incorrect data points correctly anticipated[34]. The mean TP, TN, FP, and FN were determined for every performance assessment using a following Equation: (6), (7), (8), (9), and (10), respectively. If the model properly predicts the DR positive class, then it is called TP. Furthermore, TN is capable of accurately predicting the negative class. Though it anticipates a good outcome, FP is a negative class. When the model's predictions are not based on DR, but the actual picture is DR, FN happens.

1) Accuracy

The accuracy is a proportion of the total value that is correct when considering both the positive and negative categories. The accuracy is mathematically modeled as below Equation (6):

$$Accuracy = \frac{correct\ predictions}{total\ predictions} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

2) Precision

A proportion of positive samples that were classified as TruePositives is known as precision. The mathematical format is shown below in Equation (7):

$$Precision = \frac{TP}{FP+TP} \tag{7}$$

3) Recall/Sensitivity

Recall is also called sensitivity. The recall metric measures an accuracy of a detection system by dividing a yield of positive samples by the number of genuine positives, as shown below Equation (8):

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

4) F1-Score

A single-measure score that combines the recall and precision matrices, the F1-score is an evaluation matrix that calculates the harmonic mean of the two matrices. The precision mathematical formula is shown behind Equation (9):

$$F1 - Score = 2\frac{precision \times recall}{precision + recall} \tag{9}$$

5) Specificity

For medical diagnostic purposes, a test's specificity is defined as its capacity to accurately identify healthy patients, as calculated in Equation (10):

$$Specificity = \frac{TN}{TN+FP} \tag{10}$$

6) ROC Curve

To assess how well a binary classification model is doing, one may look at the ROC curve, which plots the TPR against the FPR at different threshold levels. Better model performance is indicated by a greater AUC; an ideal model has an AUC around 1. The ROC curve is particularly useful for balanced datasets.

7) Precision-Recall Curve

Particularly with unbalanced datasets, the Precision-Recall Curve highlights the compromise between the two. The accuracy of positive predictions is measured by precision, while the model's capacity to properly identify all positive events is assessed by recall. This curve is crucial when the positive class is underrepresented, providing a better evaluation of model performance in such cases.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

C. Results of Adaptive Hybrid Net-CNN Model for Fundas Image Classification

This section provides the experimental outcomes of both binary and multiclass classification with separate sections in terms of performance measures.

1) Results of Binary Class Classification

Table 2 offers the outcomes of the proposed Adaptive Hybrid Net-CNN model on the binary class classification that gets 98.45% accuracy on APTOS 2019 data.

Table 2. Metrix performance on binary classification

Measures	Performance (%)
Accuracy	0.9845
F1-score	0.98
Recall	0.98
Precision	0.98
Sensitivity	0.98
Specificity	0.98

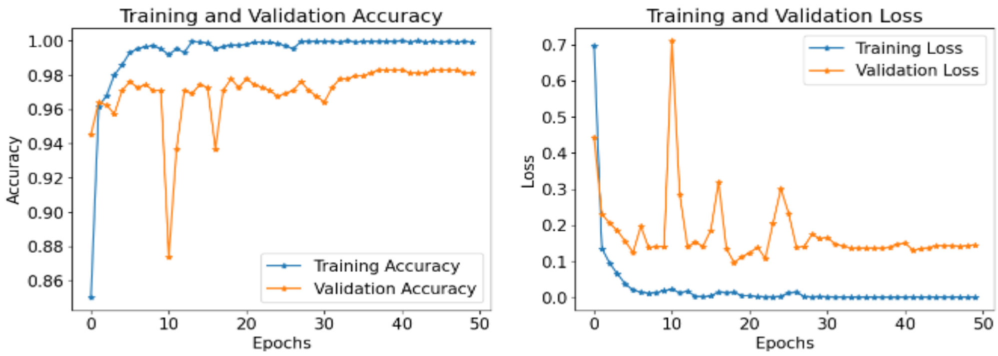


Figure 4. Accuracy/Loss plot of proposed model on binary classification

The accuracy/loss charts for training and validation of a binary classification model are displayed in Figure 4. The training accuracy reaches a high of 0.9994 with a low loss of 6.7277e-04. However, the validation accuracy plateaus around 0.9845 with a higher loss of 0.1210. When a model works well on known data but struggles on unknown data, it might be a sign of overfitting. Another sign of possible overfitting is the validation loss, which exhibits a little rising trend towards the conclusion.

	precision	recall	f1-score	support
0	0.98	0.98	0.98	275
1	0.98	0.98	0.98	311
accuracy			0.98	586
macro avg	0.98	0.98	0.98	586
weighted avg	0.98	0.98	0.98	586

Figure 5. Classification report of proposed model on binary classification

Figure 5 shows the outcomes of the prescribed model's binary classification performance in the classification report. When a model's F1score, recall, and precision are all high for both the 0 and 1 classes, it means it can accurately detect TP while minimizing FP and FN. The overall accuracy is also high at 0.98 with support 586, demonstrating a model's effectiveness in classifying samples correctly.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

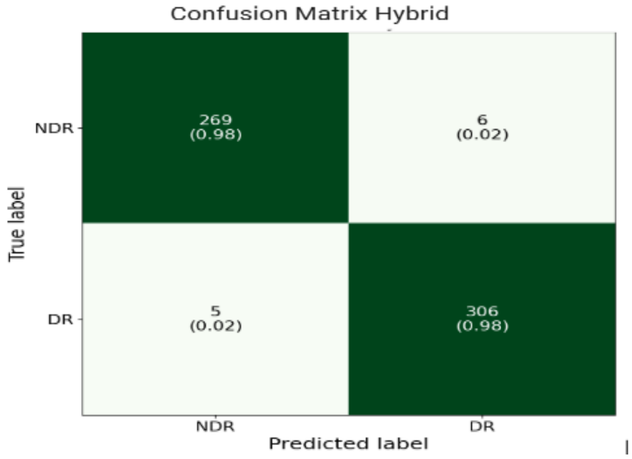


Figure 6. Confusion matrix of proposed model on binary classification

A thorough examination of the model's performance in binary classification is shown by the confusion matrix in Figure 6. According to the matrix, out of 311 DR samples, 306 were accurately identified, and 269 out of 275 NDR samples were correctly classified. The high values on the diagonal indicate that the model is accurately predicting both classes. The off-diagonal values are relatively low, suggesting that the model is effectively minimizing FP and FN.

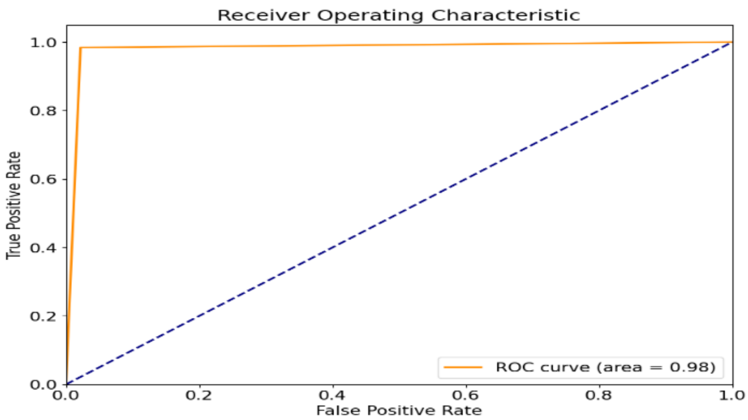
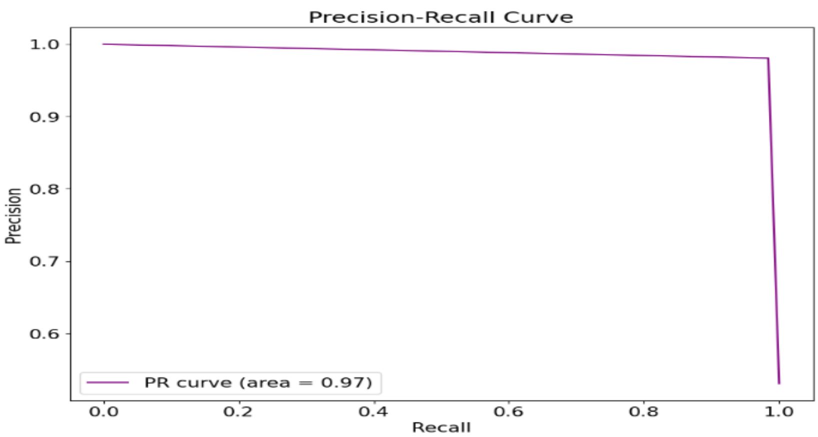


Figure 7. ROC curve of proposed model on binary classification

A suggested model's performance on binary classification is shown by the ROC curve in Figure 7. A fact that the TPR and FPR are clearly separated on the curve demonstrates a model's ability to distinguish among the two groups. With an AUC of 0.98, which is quite near to 1.0, the model seems to have excellent accuracy in properly identifying data.



Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Figure 8. PR curve of proposed model on binary classification

Figure 8 displays the Precision-Recall curve, which illustrates the correlation between the two metrics for the suggested model using binary classification. The model seems to strike a decent mix between recall and precision, as shown by the high AUC of 0.97.

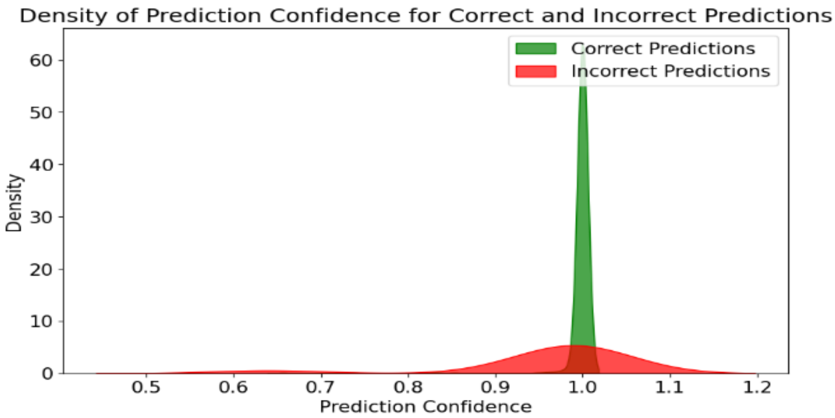


Figure 9. Density of Prediction Confidence on Binary Classification

The density plot in Figure 9 shows the distribution of prediction confidence scores for both correct and incorrect predictions of the proposed binary classification model. The plot reveals that the correct predictions tend to have higher confidence scores, clustering around 1.0, while the incorrect predictions are more spread out and have lower confidence scores. This indicates that the model is more confident in its correct predictions and less confident in its incorrect ones.

2) Results of Multi-class classification

Table 3 provides the outcomes of a proposed Adaptive Hybrid Net-CNN model on Multi-class classification that gets 96.36% accuracy on APTOS 2019 data.

Table 3. Metrix performance on multi-classification

Measures	Performance (%)
Accuracy	0.9636
F1-score	0.97
Recall	0.97
Precision	0.97
Sensitivity	1.00
Specificity	0.99

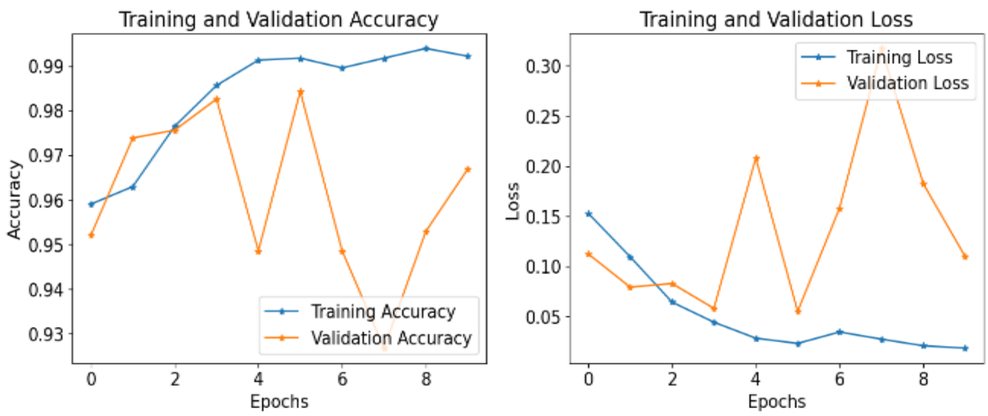


Figure 10. Accuracy/Loss plot of proposed model on Multi classification

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

The suggested model's loss curves and training and validation accuracy on a multi-classification task are displayed in Figure 10. The training accuracy steadily increases across epochs, reaching a high of 0.9928, while the training loss correspondingly decreases to 0.0177. The validation accuracy 0.9636 also shows an upward trend, indicating good generalization, although it plateaus around epoch 4. The validation loss 0.1497 exhibits a similar pattern, initially decreasing sharply and then stabilizing.

	precision	recall	f1-score	support
0	0.99	1.00	1.00	400
1	0.97	0.93	0.95	370
2	0.94	0.97	0.95	377
accuracy			0.97	1147
macro avg	0.97	0.97	0.97	1147
weighted avg	0.97	0.97	0.97	1147

Figure 11. Classification report of proposed model on Multi classification

The suggested model's classification report on a multi-classification job is shown in Figure 11. The model performs well overall, as seen by its accuracy of 0.97 and support of 1157. Additionally, the model's excellent overall performance is shown in the weighted and macro averages of these measures.

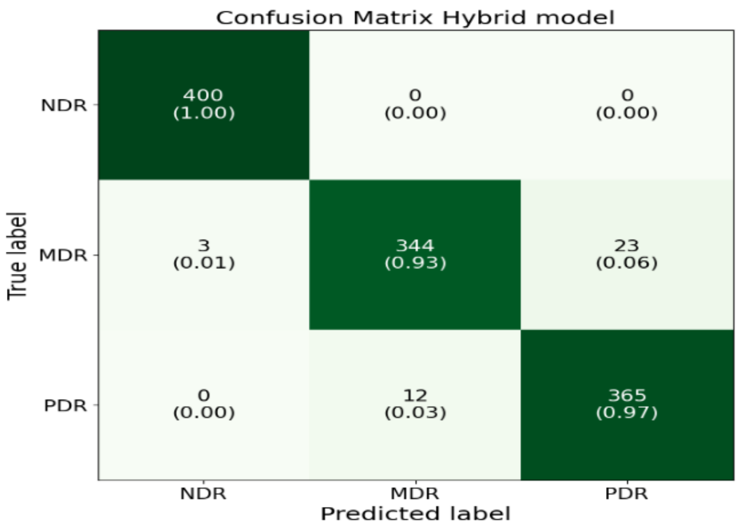


Figure 12. Confusion matrix of proposed model on Multi classification

A suggested model's confusion matrix on a multi-classification problem is displayed in Figure 12. A matrix highlights a model's performance in forecasting every class. A diagonal value (400, 344, 365) represents the number of correctly classified instances for every class (NDR, MDR, PDR, respectively). The off-diagonal values indicate misclassifications. For instance, 23 MDR instances were misclassified as PDR. The values in parentheses represent the proportion of correctly classified instances within each class, indicating high accuracy for all three classes.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

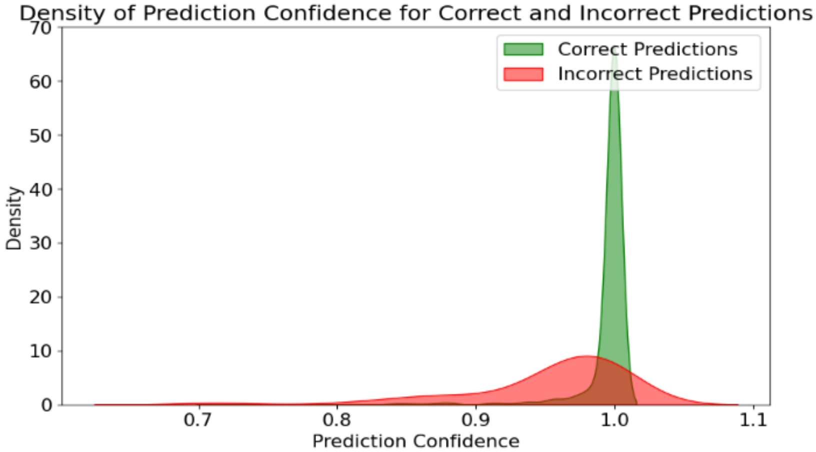


Figure 13. Density of Prediction Confidence on Multi-classification

Figure 13 visualizes the distribution of prediction confidence scores for both correct and incorrect predictions. The density of forecasts inside each confidence range is shown on the y-axis, while the prediction confidence ranges from 0.7 to 1.1 on the x-axis. The density of accurate forecasts is shown by the green curve, while the density of inaccurate predictions is shown by the red curve. The figure shows that correct predictions tend to have higher confidence scores, with a peak around 1.0, while incorrect predictions are more evenly distributed across lower confidence scores. This suggests that the model is more confident in its correct predictions than its incorrect ones.

D. Discussion

The performance comparison between various models for binary classification on the APTOS 2019 dataset reveals the superior efficacy of the Adaptive Hybrid Net-CNN, shown in Table 4. This model achieves the highest accuracy of 98.45%, significantly outperforming GoogleNet_SVM (84.85%), Hybrid_SVM (85.40%), and ResNet-18_SVM (91.26%). The Adaptive Hybrid Net-CNN also excels according to the F1score (0.98), which indicates a balanced performance between precision and recall, while the competing models record scores of 84.85%, 85.27%, and 91.26% for Google Net_SVM, Hybrid_SVM, and ResNet-18_SVM, respectively. Regarding recall, Adaptive Hybrid Net-CNN maintains its dominance with 0.98, compared to 84.85%, 85.40%, and 91.26% for the other models. Similarly, in precision, Adaptive Hybrid Net-CNN achieves 0.98, outperforming GoogleNet_SVM (84.85%), Hybrid_SVM (86.20%), and ResNet-18_SVM (91.30%). These results emphasize the robustness and effectiveness of Adaptive Hybrid Net-CNN in fundus image classification tasks, demonstrating its superior ability to accurately identify and classify retinal images in comparison to other models.

Table 4. Comparison between base and proposed models for binary classification

Measures	Adaptive Hybrid Net-CNN	Google Net_SVM	Hybrid_SVM	ResNet-18_SVM
Accuracy	0.9845	84.85	85.40	91.26
F1-score	0.98	84.85	85.27	91.26
Recall	0.98	84.85	85.40	91.26
Precision	0.98	84.85	86.20	91.30

The multi-classification performance comparison on the APTOS 2019 dataset highlights the significant advantage of the Adaptive Hybrid Net-CNN model over its counterparts, as displayed in Table 5. It achieves the highest acc-uracy of 96.36%, far exceeding Google Net_SVM (57.33%), Hybrid_SVM (46.41%), and ResNet-18_SVM (65.52%). Adaptive Hybrid Net-CNN also stands out with an F1-score of 0.97, compared to 58.64%, 44.57%, and 66.19% for Google Net_SVM, Hybrid_SVM, and ResNet-18_SVM, respectively. In terms of recall, Adaptive Hybrid Net-CNN maintains a consistent performance of 0.97, outperforming Google Net_SVM (57.33%), Hybrid_SVM (46.41%), and ResNet-18_SVM (65.52%). The model's precision is equally impressive at 0.97, while the other models achieve 60.97%, 47.01%, and 67.25% for GoogleNet_SVM, Hybrid_SVM, and ResNet-18_SVM, respectively. The efficiency of the Adaptive Hybrid Net-CNN in properly identifying retinal diseases is confirmed by these findings, which highlight its superior capabilities in handling difficult multi-class fundus image classification tasks and its resilience.

Table 5. Comparison between base and proposed models for Multi classification

Measures	Adaptive Hybrid Net-CNN	GoogleNet_SVM	Hybrid_SVM	ResNet-18_SVM
Accuracy	0.9636	57.33	46.41	65.52
F1-score	0.97	58.64	44.57	66.19
Recall	0.97	57.33	46.41	65.52
Precision	0.97	60.97	47.01	67.25

The proposed Adaptive Hybrid Net-CNN model demonstrates significant advantages in eye disease detection for both binary and multi-class classification. By integrating advanced architectures like InceptionV3, MobileNetV3Small, ResNet101, and adaptive CNN layers, the model effectively extracts robust features, leading to superior classification accuracy. Its preprocessing pipeline ensures high-quality input data, and the use of techniques like SMOTE addresses class imbalance issues. When compared to the other models, Adaptive Hybrid Net-CNN performs better across the board, including accuracy (98.45% for binary and 96.36% for multi-class classification), recall, precision, and F1score. Results from comparing the suggested model to other approaches show that it is capable of accurately diagnosing eye illnesses, with greater recall, precision, acc-uracy, and F1score, respectively. The nature of this architecture, preprocessing, and optimization enables Adaptive Hybrid Net-CNN to achieve both high precision and feasibility for implementation in real-world diagnostic tasks in clinics.

CONCLUSION AND FUTURE SCOPE

This work presents the ADCY-Dnee learning model, which uses fundus images to identify and classify eye diseases. It uses Adaptive Hybrid Net-CNN, which combines InceptionV3, MobileNetV3Small, ResNet-101, and adaptive CNN layers. The model has demonstrated significant success in both binary and multi-class classification. The model achieved near-perfect performance in the binary classification task, with an F1score of 0.98, recall of 0.98, and precision of 0.98. Its accuracy was 99.94% on the training set and 98.45% on the testing set. For the multi-class detection task, the training acc-uracy was 99.28% and the testing acc-uracy was 96.36%, with the F1 score equal to 0.97, as well as excellent class-specific sensitivity and specificity results. According to the group, these findings hold promise for the practical implementation of the suggested technique in the identification and diagnosis of ocular disorders, which can improve the clinical diagnostic and patient care capacities.

Nonetheless, the proposed methodology for eye disease classification has certain limitations despite a high accuracy rate. A significant disadvantage of the study is that the trials were carried out solely on the APTOS 2019 dataset. Therefore, any generalisation to other datasets or patient groups may be biased to some extent. Further, the model may have limitations such as the amount of light in the scene, better magnification of the image and so on. Future work might extend the dataset with distressing examples of eye diseases, optimize data augmentation techniques for different image conditions, and integrate explainable AI techniques to increase the model's interpretability. In addition, considering its real-world application in clinical practice, and how it could be incorporated with other diagnostic mechanisms can open opportunities for the model in healthcare.

REFRENCES

[1] J. Refardt, “Diagnosis and differential diagnosis of diabetes insipidus: Update,” *Best Pract. Res. Clin. Endocrinol. Metab.*, vol. 34, no. 5, p. 101398, Sep. 2020, doi: 10.1016/j.beem.2020.101398.

[2] T. L. Elliott and K. M. Pfotenhauer, “Classification and Diagnosis of Diabetes,” *Primary Care - Clinics in Office Practice*. 2022. doi: 10.1016/j.pop.2021.11.011.

[3] P. Saeedi *et al.*, “Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition,” *Diabetes Res. Clin. Pract.*, 2019, doi: 10.1016/j.diabres.2019.107843.

[4] A. Pandey, S. Chawla, and P. Guchhait, “Type-2 diabetes: Current understanding and future perspectives,” *IUBMB Life*, 2015, doi: 10.1002/iub.1396.

[5] K. G. M. M. Alberti, P. Zimmet, and J. Shaw, “Metabolic syndrome - A new world-wide definition. A consensus

statement from the International Diabetes Federation," *Diabetic Medicine*. 2006. doi: 10.1111/j.1464-5491.2006.01858.x.
[6] C. Mohanty *et al.*, "Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy," *Sensors*, 2023, doi: 10.3390/s23125726.
[7] R. E. Hacisoftaoglu, M. Karakaya, and A. B. Sallam, "Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems," *Pattern Recognit. Lett.*, 2020, doi: 10.1016/j.patrec.2020.04.009.
[8] D. Mane, R. Ashtagi, R. Jotrao, P. Bhise, P. Shinde, and P. Kadam, "Diabetic Retinopathy Detection using Deep Learning," *J. Electr. Syst.*, 2023, doi: 10.52783/jes.687.
[9] A. Saini, K. Guleria, and S. Sharma, "An Efficient Deep Learning Model for Eye Disease Classification," in *Proceedings - International Research Conference on Smart Computing and Systems Engineering, SCSE 2023*, 2023. doi: 10.1109/SCSE59836.2023.10215000.
[10] N. Asiri, M. Hussain, F. Al Adel, and N. Alzaidi, "Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey," *Artificial Intelligence in Medicine*. 2019. doi: 10.1016/j.artmed.2019.07.009.
[11] J. W. Fan *et al.*, "Applications of Deep Learning Techniques for Diabetic Retinal Diagnosis," *Zidonghua Xuebao/Acta Automatica Sinica*. 2021. doi: 10.16383/j.aas.c190069.
[12] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, 2016, doi: 10.1016/j.neucom.2015.09.116.
[13] R. Raman, S. Srinivasan, S. Virmani, S. Sivaprasad, C. Rao, and R. Rajalakshmi, "Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy," *Eye (Basingstoke)*. 2019. doi: 10.1038/s41433-018-0269-y.
[14] R. Vega, G. Sanchez-Ante, L. E. Falcon-Morales, H. Sossa, and E. Guevara, "Retinal vessel extraction using Lattice Neural Networks with dendritic processing," *Comput. Biol. Med.*, 2015, doi: 10.1016/j.compbiomed.2014.12.016.
[15] N. Sikder, M. Masud, A. K. Bairagi, A. S. M. Arif, A. Al Nahid, and H. A. Alhumyani, "Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images," *Symmetry (Basel).*, 2021, doi: 10.3390/sym13040670.
[16] N. Dhungel, G. Carneiro, and A. P. Bradley, *Deep Learning and Convolutional Neural Networks for Medical Image Computing*. 2017.
[17] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*. 2017. doi: 10.1016/j.media.2017.07.005.
[18] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015. doi: 10.1109/CVPR.2015.7298594.
[19] G. Zeng, Y. He, Z. Yu, X. Yang, R. Yang, and L. Zhang, "InceptionNet/GoogLeNet - Going Deeper with Convolutions," *Cvpr*, 2016.
[20] M. Aamir *et al.*, "An Adoptive Threshold-Based Multi-Level Deep Convolutional Neural Network for Glaucoma Eye Disease Detection and Classification," *Diagnostics*, 2020, doi: 10.3390/diagnostics10080602.
[21] G. Ruminaviciute and J. Bernataviciene, "Ensemble for quality assessment of eye fundus images involving deep learning methods," in *Iberian Conference on Information Systems and Technologies, CISTI*, 2023. doi: 10.23919/CISTI58278.2023.10211416.
[22] M. A. Abini and S. S. S. Priya, "A Deep Learning Framework for Detection and Classification of Diabetic Retinopathy in Fundus Images Using Residual Neural Networks," in *9th International Conference on Smart Computing and Communications: Intelligent Technologies and Applications, ICSCC 2023*, 2023. doi: 10.1109/ICSCC59169.2023.10335079.
[23] R. Chavan and D. Pete, "Classification of Retinal Fundus Images using VGG16 and Inception V3," in *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*, 2023. doi: 10.1109/ICCCNT56998.2023.10306593.
[24] H. Sharma, J. Wasim, and P. Sharma, "Analysis of Eye Disease classification by Fundus Images using Different Machine/Deep/Transfer Learning Techniques," in *2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM)*, 2024, pp. 1–6. doi: 10.1109/ICIPTM59628.2024.10563429.

[25] M. Oulhadj *et al.*, “Diabetic Retinopathy Prediction Based on a Hybrid Deep Learning Approach,” in *2024 International Conference on Intelligent Systems and Computer Vision (ISCV)*, 2024, pp. 1–5. doi: 10.1109/ISCV60512.2024.10620125.

[26] P. Uttakit, P. Hanpinitsak, S. Thanapaisal, and P. Polpinit, “Classification of Glaucoma Severity Stages Using Deep Learning,” in *2024 21st International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2024, pp. 1–5. doi: 10.1109/ECTI-CON60892.2024.10594822.

[27] T. Malik, V. Nandal, and P. Garg, “Deep Learning-Based Classification of Diabetic Retinopathy: Leveraging the Power of VGG-19,” in *2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP)*, 2024, pp. 645–651. doi: 10.1109/INNOCOMP63224.2024.00111.

[28] S. Dutta, B. C. S. Manideep, S. M. Basha, R. D. Caytiles, and N. C. S. N. Iyengar, “Classification of diabetic retinopathy images by using deep learning models,” *Int. J. Grid Distrib. Comput.*, 2018, doi: 10.14257/ijgdc.2018.11.1.09.

[29] B. R.M., K. B. Vardhan, M. Nidhish, S. Kiran C., D. Nahid Shameem, and V. Sai Charan, “Eye Disease Detection Using Deep Learning Models with Transfer Learning Techniques,” *ICST Trans. Scalable Inf. Syst.*, vol. 11, pp. 1–13, 2024, doi: 10.4108/eetsis.5971.

[30] D. Pang, Z. Guan, T. Luo, W. Su, and R. Dou, “Real-time detection of road manhole covers with a deep learning model,” *Sci. Rep.*, 2023, doi: 10.1038/s41598-023-43173-z.

[31] J. Z. Wang *et al.*, “Classification of Color Fundus Photographs Using Fusion Extracted Features and Customized CNN Models,” *Healthc.*, 2023, doi: 10.3390/healthcare11152228.

[32] M. Byra, “Breast mass classification with transfer learning based on scaling of deep representations,” *Biomed. Signal Process. Control*, 2021, doi: 10.1016/j.bspc.2021.102828.

[33] N. Minh and K. Nguyen, “Real-time fashion items classification using TensorflowJS and ZalandoMNIST dataset,” no. August, 2019.

[34] F. M. J. Mehedi Shamrat *et al.*, “An advanced deep neural network for fundus image analysis and enhancing diabetic retinopathy detection,” *Healthc. Anal.*, 2024, doi: 10.1016/j.health.2024.100303.