

Fundus image classification using feature concatenation for early diagnosis of retinal disease

DIGITAL HEALTH

Volume 11: 1–24

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/20552076251328120

journals.sagepub.com/home/dhj



Sara Ejaz¹, Hafiz U Zia¹, Fiaz Majeed¹, Umair Shafique¹,
Stefania Carvajal Altamiranda^{2,3,4}, Vivian Lipari^{2,5,6} and
Imran Ashraf⁷ 

Abstract

Background: Deep learning models assist ophthalmologists in early detection of diseases from retinal images and timely treatment.

Aim: Owing to robust and accurate results from deep learning models, we aim to use convolutional neural network (CNN) to provide a non-invasive method for early detection of eye diseases.

Methodology: We used a hybridized CNN with deep learning (DL) based on two separate CNN blocks, to identify multiple Optic Disc Cupping, Diabetic Retinopathy, Media Haze, and Healthy images. We used the RfMiD dataset, which contains various categories of fundus images representing different eye diseases. Data augmenting, resizing, coping, and one-hot encoding are used among other preprocessing techniques to improve the performance of the proposed model. Color fundus images have been analyzed by CNNs to extract relevant features. Two CNN models that extract deep features are trained in parallel. To obtain more noticeable features, the gathered features are further fused utilizing the Canonical Correlation Analysis fusion approach. To assess the effectiveness, we employed eight classification algorithms: Gradient boosting, support vector machines, voting ensemble, medium KNN, Naive Bayes, COARSE- KNN, random forest, and fine KNN.

Results: With the greatest accuracy of 93.39%, the ensemble learning performed better than the other algorithms.

Conclusion: The accuracy rates suggest that the deep learning model has learned to distinguish between different eye disease categories and healthy images effectively. It contributes to the field of eye disease detection through the analysis of color fundus images by providing a reliable and efficient diagnostic system.

Keywords

Public health, retinal disease detection, deep learning, feature extraction, convolutional neural networks

Received: 23 July 2024; accepted: 4 February 2025

Introduction

The human retina, a delicate and intricate neural tissue lining the back of the eye, plays a pivotal role in vision by

grasping and interpreting light signals. However, this vital organ is susceptible to various pathologies that, if undetected and untreated, can lead to irreversible vision impairment. Among these, diabetic retinopathy (DR),

¹Department of Information Technology, University of Gujrat, Gujrat, Pakistan

²Universidad Europea del Atlantico, Santander, Spain

³Universidade Internacional do Cuanza, Cuito, Bie, Angola

⁴Fundacion Universitaria Internacional de Colombia, Bogota, Colombia

⁵Universidad Internacional Iberoamericana, Campeche, Mexico

⁶Universidad de La Romana, La Romana, Republica Dominicana

⁷Department of Information and Communication Engineering, Yeungnam University, Gyeongsan South Korea

Corresponding author:

Imran Ashraf, Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, South Korea.
Email: imranashraf@ynu.ac.kr



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

media haze (MH), and optic disc cupping (ODC) stand out as significant contributors to visual morbidity.

Globally, approximately 2.2 billion individuals experience either near or distance vision impairments. In at least 1 billion of these cases, vision impairment is either preventable or has not yet been addressed.¹ DR, a common complication of diabetes, poses a substantial threat to vision health. The condition manifests as damage to the blood vessels in the retina, leading to leakage and abnormal growth.

DR has the potential to result in various severe eye complications. Over time, approximately 1 in 15 individuals with diabetes will experience diabetic macular edema (DME). DME occurs when blood vessels in the retina release fluid into the macula, a vital region of the retina essential for clear central vision, leading to blurred vision. The progressive and irreversible loss of sight caused by glaucoma is caused by increased pressure within the eyeball.² Additionally, DR can stimulate the growth of abnormal blood vessels that extend from the retina, obstructing the drainage of fluid from the eye and giving rise to a form of glaucoma, a group of eye diseases known for causing visual impairment and blindness.³ The identification of DR through automated means is crucial, given that it stands as the primary contributor to irreversible vision impairment among the working-age populace in developed nations.⁴ Figure 1 shows the normal human eye with various parts outlined.

MH, characterized by opacities in the ocular media, can impede the clarity of fundus images, complicating the diagnosis of underlying retinal conditions. The opacity in the eye caused by MH can serve as an indicator for the onset of conditions such as cataracts, corneal swelling, vitreous cloudiness, or constricted pupils. Therefore, it is crucial to promptly and precisely diagnose MH to prevent potential vision loss that may result if the condition is not treated in a timely manner.⁵

ODC, often associated with glaucoma, involves the excavation of the optic nerve head, causing progressive damage to the optic nerve fibers. In humans, glaucoma is the second leading cause of blindness, and the number of cases is steadily increasing.⁶⁻⁸ Optic nerve cupping is divided into two primary categories. The initial classification is attributed to injury or trauma, whereas the second arises from diverse medical conditions or diseases.⁹ Several medical conditions associated with optic nerve cupping encompass

- Optic nerve head drusen
- Glaucoma
- Optic nerve atrophy
- Optic neuritis

The traditional methods of diagnosing retinal diseases have been labor-intensive, and reliant on manual examination by

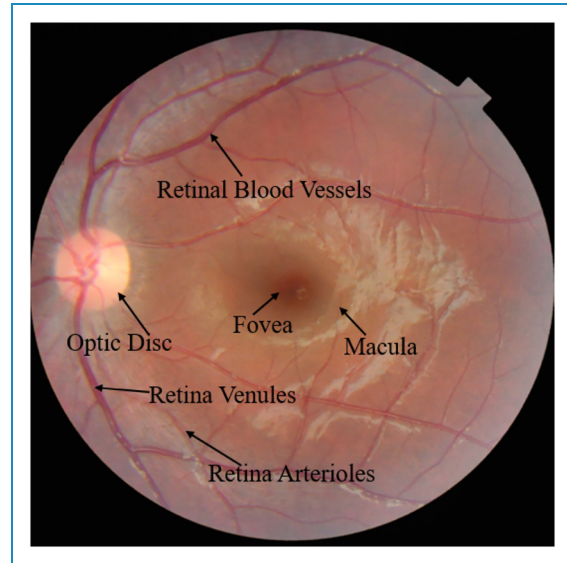


Figure 1. Normal human eye with various parts.

skilled ophthalmologists. However, with the emergence of deep learning (DL) technologies, particularly within the realm of convolutional neural networks (CNNs), there has been a paradigm shift in the approach to retinal disease diagnosis. It is promising that DL models are capable of learning intricate patterns and features automatically based on vast datasets, providing a route to accurate and efficient diagnosis.

Variability in image quality, noise, and low-contrast conditions can impair the ability of models to accurately capture disease-specific features, often limiting their performance in real-world applications. Current methodologies for retinal disease classification primarily rely on single CNN architectures, which, while effective in some cases, may lack robustness across diverse datasets. Existing approaches can struggle with feature extraction, particularly in noisy or low-contrast images, resulting in inconsistent accuracy. Furthermore, models that do not effectively integrate multi-source features may fail to capture the complexity of retinal diseases, reducing their ability to generalize across cases.

To address these challenges, we propose a novel methodology utilizing dual CNNs for feature extraction, followed by canonical correlation analysis (CCA) fusion to enhance feature representation. By employing two CNNs, the proposed model captures a broader range of characteristics within retinal images, and CCA fusion effectively combines these complementary features, creating a comprehensive representation. This study initially focuses on evaluating the efficacy of raw feature extraction and CCA fusion in classifying retinal disease without the use of enhancement techniques such as contrast adjustment or noise handling.

This foundational approach allows for a baseline assessment of model performance under standard imaging

conditions, providing a basis for future integration of contrast enhancement and noise-handling methods. Ultimately, the methodology aims to improve the robustness and accuracy of retinal disease classification, contributing to the development of more reliable diagnostic tools for clinical applications.

Research objective

The following are the research objectives of the current study

- To design and implement a hybrid CNN architecture for extracting suitable features from retinal fundus.
- To apply CCA for fusing features extracted by the hybrid CNN. This fusion aims at capturing complementary information from multiple feature spaces.
- To make use of benchmark datasets RFMiD and RFMiD 2.0 for model training and validation. Testing is to analyze model's robustness and generalizability across different datasets.
- To evaluate the proposed model against existing state-of-the-art models to demonstrate its superior performance in terms of classification accuracy.
- To contribute to automated medical diagnosis by providing a more accurate and reliable tool for the classification of retinal diseases, aiding ophthalmologists in early detection and treatment.

Research contributions

The following key contributions are made in this study.

- A novel hybrid CNN architecture is introduced that improves feature extraction from retinal fundus images by capturing low-level and high-level features more effectively than traditional CNN models.
- The CCA fusion method is utilized in this study for feature fusion. Features extracted using multiple CNN models are combined for enhanced feature representation and improved classification performance.
- The proposed hybrid CNN model is used for experiments on the latest RFMiD and RFMiD 2.0 datasets for performance evaluation in comparison to existing state-of-the-art approaches. A thorough evaluation of the model is carried out concerning robustness and ability to generalize across different datasets.
- The proposed model enhances the tools available for the automated classification of retinal diseases, potentially improving clinical workflows and patient outcomes by aiding in the early and accurate diagnosis of retinal conditions.

The rest sections of the article include the following. Section "Overview of existing literature" provides the summary of related work. Section "Materials and methods" presents the suggested method, covering aspects such as image acquisition, preprocessing, data augmentation, feature extraction, and classification. In Section "Results", the statistics obtained post-training for the customized CNN models are outlined and classified using machine learning models. The study is concluded in Section "Conclusion and future direction".

Overview of existing literature

In literature, retina disease classification using DL has been extensively explored to improve diagnostic accuracy and efficiency. Numerous studies have demonstrated the efficacy of identifying and differentiating between various retinal conditions.¹⁰

In study by Das et al.¹¹ the author developed and evaluated a compact CNN using four retinal image datasets: DRD,¹² Messidor-2,¹³ IDRiD,¹⁴ and RFMiD,¹⁵ Employing a 12-fold cross-validation technique, the model achieved notable accuracy: 79.96% on DRD, 94.75% on Messidor-2, 96.74% on IDRiD, and 89.10% on RFMiD. These results highlight the model's effectiveness and adaptability across various datasets, providing a valuable tool for the early detection of retinal diseases and improving patient care in ophthalmology. The study¹⁶ explored the use of various DL models to detect eye diseases using fundus imaging. An automated system was developed to process and enhance a dataset of 4697 images through brightness and contrast adjustments, feature extraction, data augmentation, and image classification using CNN. Among the five models evaluated, ResNet152 proved the most effective, achieving an AUC score of 96.47%. The paper also includes visualizations of model predictions, highlighting confidence scores and heatmaps that indicate focal points, especially where lesions are detected.

The study by Nagamani and Rayachoti¹⁷ aims to develop a DL model using OCT images to enhance the classification and segmentation of retinal diseases. It classifies volumetric OCT images, recognizing conditions such as DME, CNV, AMD, and DN. The research introduces a Modified ResNet-50 approach and uses a Bi-LSTM-based deep recurrent CNN for image segmentation. The model, tested on publicly available datasets, achieves 99.76%. Similarly,¹⁸ introduces the Fundus-DeepNet system, an automated multi-label DL classification system for identifying multiple ocular diseases in fundus images. A comprehensive pre-processing process involves cropping, resizing, enhancing contrast, removing noise, and enhancing data. Deep feature representations are then extracted using High-Resolution Network and Attention Block, further enhanced by SENet Block to consolidate them into a single representation. Finally, a discriminative restricted

Boltzmann machine classification model, incorporating a Softmax layer, generates a probability distribution for identifying eight ocular diseases.

The authors address the challenge of automatic detection of retinal diseases, emphasizing the limitations associated with low contrast, illumination inhomogeneity, convergence rates, overfitting, and classification errors.¹⁹ The proposed approach employs ensemble-based DL techniques for enhanced retinal disease prediction, structured into pre-processing, adaptive Gaussian kernel PDF-based matched filtering, and post-processing for segmentation, and classification stages. The classification stage employs three approaches, EfficientNet B0, VGG16, and ResNet-152²⁰ where the feature vectors are fused through an ensemble approach. The proposed method demonstrates impressive performance metrics, including 99.71% accuracy, 98.63% precision, 98.25% recall, and 99.22% F-measure.

The authors²¹ introduce an automated DL-based framework for the non-invasive diagnosis of multiple eye diseases using color fundus images. The study utilizes a multi-class eye disease dataset RFMiD¹⁵ to develop an efficient diagnostic system. The framework involves extracting multi-class fundus images from a multi-label dataset and applying various augmentation techniques to enhance real-time robustness. A multi-layer neural network is developed to train and test images for the diagnosis of different eye problems. The key component extracts relevant features from the input color fundus image dataset, and these processed features are used for prognostic diagnostic determinations.

The study by Pan et al²² aims to enhance ophthalmic diagnostics through an automated deep-learning system. In total, 1032 fundus images were gathered from 516 patients using a fundus camera. InceptionV3 and ResNet-50 DL models were employed for classification. A 93.81% accuracy was achieved for ResNet-50 and 91.76% accuracy was achieved for Inception V3. The research serves as a reference for clinical diagnosis or screening of DR and other eye diseases. Along the same directions,²³ addresses the challenge of automatically detecting disease states of the retina by developing a model VGG-19²⁴ architecture. The model utilizes transfer learning and is educated on an extensive dataset comprising 84,568 cases of OCT²⁵ retinal images, covering four conditions: CNV, DN, DME, and normal retinal form. The proposed model achieves a remarkable classification accuracy of 99.17%, with specificities of 0.995 and sensitivity of 0.99, surpassing existing models.

The research by Pandey et al.²⁶ endeavors to develop an algorithm for classifying various retinal pathologies in fundus photographs. The researchers utilized a deep convolutional ensemble comprising five CNNs to categorize retinal images into DR, glaucoma, ARMD, and normal eyes. The CNN architecture was based on the InceptionV3 model with pretraining on the ImageNet dataset, using 43,055 images from 12 datasets. They used DiaretDB,²⁷

Drishti-GS,²⁸ DRIVE,²⁹ HRF,³⁰ IDRiD,¹⁴ Kaggle-39,³¹ Kaggle-DR, ODIR,³² MESSDIDOR,³³ ORIGA-light,³⁴ REFUGE,³⁵ and STARE³⁶ datasets. The study found that the DCE achieved a mean accuracy of 79.2%.

The research by Thanki et al.³⁷ introduces an innovative computer-aided triage system that incorporates a DL and ML for the development and analysis of color retinal fundus images, specifically aimed at classifying images indicative of glaucoma. The methodology involves extracting deep features from retinal images through a deep neural network, then classification and analysis utilizing various ML classifiers. The experimental findings demonstrate that the integration of a DL with a logistic regression-based classifier surpasses the performance of existing glaucomatous triage systems.

The study Almustafa et al.³⁸ employs the STARE dataset,³⁶ comprising 385 retinal images with various defects. Pre-processing techniques, including augmentation and normalization, are applied to refine features for training DL algorithms. The paper evaluates five DL models EfficientNet, 3-Layers CNN, VGG-16, InceptionV2, ResNet-50, and tuning hyperparameters such as batch size. EfficientNet emerges as the best-performing model, achieving 98.43% accuracy. For each of these 14 retinal defects, unique model configurations, hyperparameter tuning, and preprocessing techniques are credited with its success.

To enhance clinical usability, the study by Ho et al.³⁹ aimed to simultaneously detect multiple ophthalmic pathologies. The researchers utilized 2560 images from the RFMiD, dividing them into training (1920) and validation (640) sets. To predict the presence of any pathology and categorize 28 different pathologies, five CNN architectures were selected and trained. To optimize training, models were designed to minimize asymmetric loss, a modified version of binary cross-entropy. As a result of the ensemble network, an AUROC score of 0.9613 was demonstrated for disease screening. Among the individual models, the SE-ResNeXt architecture achieved the highest single network score at 0.9586.

A CNN has been effectively applied to color fundus images for automated glaucoma detection. On the ACRIMA⁴⁰ database, the method achieved high performance, with an accuracy of 96.64%, sensitivity of 96.07%, specificity of 97.39%, and precision of 97.74%.⁴¹ The research by Rodriguez et al.⁴² utilizes the MuReD dataset constructed from publicly available fundus disease classification datasets. Image data quality is enhanced through a series of processing steps that encompass a broad range of diseases. It shows improvements in AUC scores of 7.9% for disease detection and 8.1% for disease classification over state-of-the-art approaches for the same task.

The authors utilize customized particle swarm optimization (CPSO) combined with four advanced machine-learning classifiers to enhance glaucoma prediction performance. It operates through five main phases preprocessing,

segmentation, feature extraction, selection of the best-scored features, and classification using the CPSO-based classifier. The images are sourced from the publicly available Digital Retinal Images for Optic Nerve Segmentation dataset. These features are then employed for training and testing, generating multiple result sets from various CPSO and supervised machine-learning classifier combinations.⁴³

The author's focus is on the development of an intelligent algorithm using DL for the classification.⁴⁴ A dataset comprising 501 images, including normal eyes and those with RVO. Fundus disease specialists categorized the images into four groups: healthy fundus, RVO, BRVO, CRVO, and MBRVO. The ResNet18⁴⁵ network model was employed for diagnosis. The intelligent system exhibited a specificity of 100% for healthy fundus. For the RVO groups, various attention mechanisms yielded specificities ranging from 0.45 to 0.91, with the ResNet18+ model achieving the highest specificities and accuracy across all groups.

The study by Elangovan and Nath⁴⁶ introduces a deep ensemble model using stacking ensemble learning for classifying glaucomatous and normal fundus images, leveraging 13 pre-trained CNN models in 65 configurations. A two-stage ensemble selection with probability averaging and support vector machine (SVM) final classification achieves robust performance. Testing on modified databases (DRISHTI-GS1-R, ORIGA-R, RIM-ONE2-R, LAG-R, ACRIMA-R) shows accuracies of up to 99.6%.

The authors in Abitbol et al.⁴⁷ aim to evaluate the capability of a DL framework for differentiating between DR, SCR, RVOs, and normal eyes using ultra-widefield color fundus photography. The study employs cross-validation and augmentation techniques for robust performance, utilizing an Adam optimizer for training. The model achieves its best performance at 10 epochs, yielding an accuracy of 88.4%. Specific disease-wise assessments reveal notable results: For DR, an accuracy of 85.2%; for RVO, accuracy is 88.4%; for SCR, accuracy is 93.8%; and for Healthy, an accuracy of 86.2%.

The objective of Kumar and Bindu⁴⁸ fundus imaging is to examine eye-related anomalies. The framework involves preprocessing steps such as contrast enhancement, over-sampling, resizing, and normalization. Densenet201 and EfficientNetB4 are employed for disease risk detection, and ResNet105 is added for multi-disease classification. The proposed framework is trained and validated on the RFMiD and tested on the ODIR dataset. The systematic review is given in Table 1.

Materials and methods

To identify retinal disorders using fundus images, we propose DL and ML algorithms. Data is gathered from two datasets RFMiD¹⁵ and RFMiD 2.0.⁴⁹ Multi-labeled and single-labeled images are included in these datasets. Single-label diseases are separated and diseases with more images are selected. The suggested methodology is shown in Figure 2. The diseases

we select contain four classes given in Table 2. After acquiring the dataset we perform preprocessing steps. Due to the different sizes of images in the dataset, preprocessing involves resizing them to the same size, augmentation to extend and balance the dataset, then cropping the unwanted area to increase the efficacy of the model and partitioning them into training and testing sets. To reduce computing time, we convert images into an array and perform one hot encoder.

Further, we implement two CNN models to extract features for three retinal diseases and one healthy class. After feature extraction from two CNNs CCA fusion is utilized to concatenate the features and apply ML algorithms to classify retinal images and healthy class. Results using ML models are analyzed in the context of accuracy, recall, sensitivity, precision, and F1 score.

Data acquisition

The dataset was compiled from RFMiD¹⁵ and RFMiD 2.0,⁴⁹ with a focus on eye disease classes that contained over 100 fundus images. The process involved converting the originally multi-labeled dataset into a multi-class dataset, essentially transforming it from a multi-label object detection problem to a multi-class classification problem. In the final dataset, images were chosen based on their affiliation with a single class and having more than 100 images in their respective categories. Excluding the normal category, three specific diseases DR, MH, and ODC were singled out from a pool of 49 diseases. The Total Number of images with classes DR, MH, ODC, and WNL are 471, 334, 172, and 931, respectively as mentioned in Table 2.

A few sample images concerning each class from the datasets are shown in Figure 3.

Preprocessing

Preprocessing is crucial for enhancing the quality of image visualization, significantly impacting the success and accuracy of subsequent stages in the proposed method. Medical images often present additional challenges such as poor quality or extraneous content, which can hinder effective visualization. Addressing these issues is essential, as low-quality images can result in unsatisfactory outcomes.⁵⁰ By using histogram equalization techniques to enhance images, details are lost and the impression is artificial.^{51,52} In the preprocessing phase, techniques such as data augmentation, resizing, cropping, and one-hot encoding are employed to enhance model efficiency.

Resize. Upon analyzing images across all four classes DR, MH, ODC, and WNL, the images' shapes in the dataset varied. To address this, all images were synthesized to have a uniform shape of $224 \times 224 \times 3$. This standardization ensures that the model is less prone to errors. Training on smaller images offers the advantage of quicker iterations and more efficient experimentation during the development

Table 1. Review of related studies.

Year	Ref.	Dataset	Classes	Method	Results
2024	Das et al. ¹¹	DRD, Messidor, IDRiD, RFMiD	5	DL	Accuracy: RFMiD 89.10% IDRiD 96.74 % Messidor-2 94.75% DRD 79.96%
2024	Nguyen et al. ¹⁶	Kangbuk Samsung Hospital	2	DL ResNet152 Vision Transformer InceptionResNetV2 RegNet ConVNext	Accuracy ResNet152 89.17% Vision Transformer 87.26% InceptionResNetV2 88.11% RegNet 88.54% ConVNext 89.08%
2024	Pandey et al. ²⁶	RFMiD, RFMiD 2.0	4	CNN	Accuracy 88.72%
2024	Nagamani and Rayachoti ¹⁷	OCT	5	ResNet-50	Accuracy 99.76%
2024	Al-Fahdawi et al. ¹⁸	OIA-ODIR	8	Fundus-DeepNet	AUC 99.86%
2023	Kumar and Singh ¹⁹	Fundus Images	11	ResNet-152 EfficientNet B0 VGG 16	Accuracy 99.71%
2023	Sengar et al. ²¹	RFMiD	4	EyeDeep-Net	Accuracy Validation 82.13% Testing 76.04%
2023	Pan et al. ²²	Hospital Shenzhen University	3	Pre-Trained DL Model Inception V3 ResNet-50	Accuracy: ResNet 93.81% InceptionV3 91.76%
2023	Choudhary et al. ²³	OCT	4	VGG-19	Accuracy 99.17%
2023	Pandey et al. ²⁶	DiaretDB, Drishti-GS, HRF, DRIVE, IDRiD, Kaggle-39, Kaggle-DR, STARE, ODIR, MESSIDOR, ORIGA-light, REFUGE	4	Deep convolutional Ensemble	Accuracy 79.2%
2023	Thanki ³⁷	DRISTHI-GS	2	DNN (Features Extraction) ML (Classification)	Accuracy 99.6%
2023	Almustafa et al. ³⁸	STARE	15	DL ResNet-50 EfficientNet 3-Layer CNN VGG-16 InceptionV2	Accuracy ResNet-50 54.61% EfficientNet 98.43% 3-Layer CNN 80.37% VGG-16 87.50% InceptionV2 96.87%
2022	Ho et al. ³⁹	RFMiD	29	Deep Ensemble Learning	Sensitivity for all 29 class Range 0.00–1.00
2022	Rodríguez et al. ⁴²	MuReD	20	Transformer	Give different Accuracy for 20 classes ranging from 0.87–0.99
2022	Xu et al. ⁴⁴	Nanjing Medical University	4	DL ResNet-18 ResNet18+SE ResNet18+CBAM ResNet18+CA	Accuracy healthy 100% BRVO 94.64% CRVO 98.21% MBRVO 96.43%
2022	Abitbol et al. ⁴⁷	Creteil University Hospital	4	DL	Accuracy RVO 88.4% DR 85.2% CSR 93.8% Healthy 86.2%
2021	Kumar and Bindu ⁴⁸	RFMiD	29	Ensemble CNN	F1 Score 94.32%
2021	Elangovan and Nath ⁴¹	ACRIMA	2	CNN	Accuracy 96.64%

DL: deep learning; CNN: convolutional neural network.

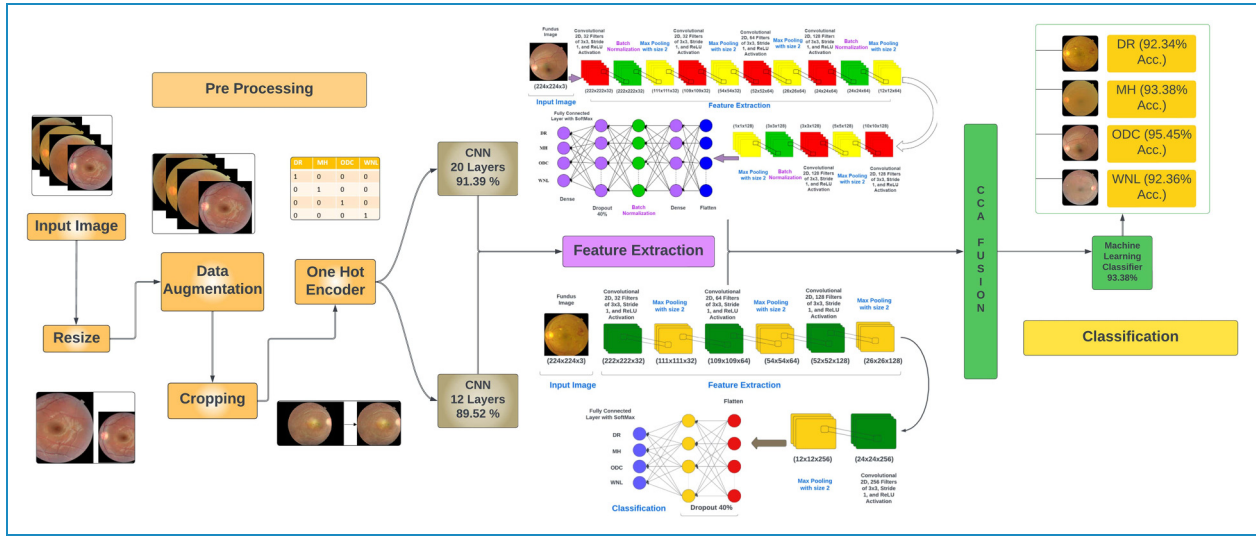


Figure 2. Proposed approach's architecture.

Table 2. Number of images in RFMiD and RFMiD 2.0 datasets.

Diseases	RFMiD	RFMiD 2.0	Total
DR	401	70	471
MH	315	19	334
ODC	155	17	172
WNL	669	262	931

DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping.

process.⁵³ This approach enables faster model training, facilitating a more streamlined and agile development process.

Data augmentation. Data augmentation is indeed a crucial step in preprocessing for machine learning tasks, especially in image classification. The Literary works^{21,54,55} use data augmentation for classification. We gather 1908 images from the dataset. We organized the dataset into four distinct classes, namely DR, MH, ODC, and normal (WNL). At the outset, the dataset includes 334 images depicting MH, 471 images depicting DR, 172 images depicting ODC, and 931 images of WNL as shown in Table 2. By applying various transformations to the original data, such as flipping, rotating, scaling, or adding noise, we can increase the variety of the dataset, thereby reducing overfitting probability and helping the model generalize better to unseen data.

Flipping an image horizontally or vertically creates mirror images while rotating an image at different angles introduces additional variations. We use horizontal flips and rotation at different angles such as 60°, 65°, 80°, 90° Sengar et al.²¹ used for fundus images. Figure 4 shows the

original image of ODC and different augmented images. After implementing data augmentation on the dataset to address the problem addressing data overfitting. Moreover, we encountered a significant class imbalance issue where the WNL class had a substantially higher number of images compared to the other classes. This created a challenge as it could potentially introduce biases in the results. To tackle this problem, we implemented data augmentation techniques to balance the classes. For classification, we split the datasets into 70:20:10 for training and testing, and validation sets. This implies that 70% of randomly selected Images were employed during the training phase, while 20% were set aside for testing and 10% was used for validation.

Cropping. Cropping as a preprocessing step is essential in many image-related tasks, helping to focus on the most relevant parts of the data, improve model performance, and reduce computational demands. It involves removing unwanted outer areas from an image or signal to focus on the region of interest. In terms of classification accuracy, cropping-based image classification performed better than non-cropping-based image classification.⁵⁶

One-hot encoder. The one-hot encoder is used to transform categorical labels into a numerical format that the ML model can interpret. By converting each class label into a binary vector with a single “1” representing the class and “0”s elsewhere, one-hot encoding enables a clear and distinct representation of each class. This transformation supports compatibility with neural network architectures by allowing the use of categorical cross-entropy as a loss function, which compares these one-hot encoded labels with model predictions. Ultimately, this technique enhances the model’s ability to differentiate between classes effectively,

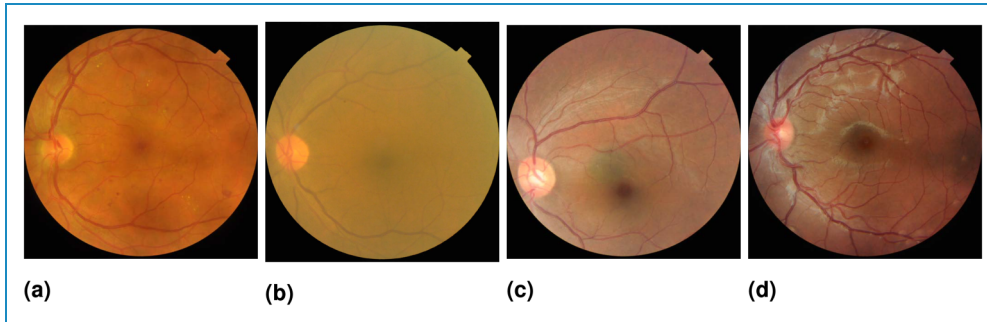


Figure 3. Sample images for, (a) Diabetic retinopathy, (b) Media haze, (c) Optic disc cupping, and (d) Healthy eye.

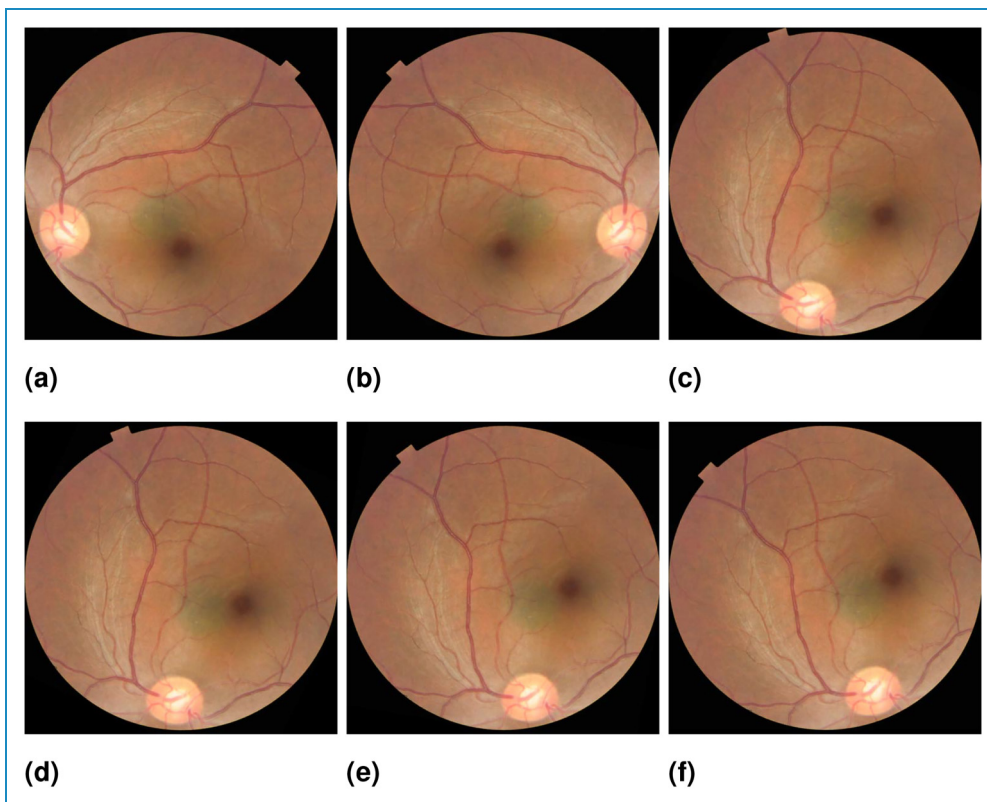


Figure 4. (a) Original (b) flipped (c) 60° rotation (d) 65° rotation (e) 80° rotation (f) 90° rotation.

improving its capacity to learn and generalize in multi-class image classification tasks.

Experimental setup. The study used Python to conduct trials on a 64-bit version of the Windows 10 operating system. This system was powered by an Intel Core i5 7th Generation CPU, contained 8 GB of RAM, and offered a 237 GB storage capacity.

Proposed deep learning architecture

DL is highly effective for image feature extraction due to its capability to autonomously acquire hierarchical patterns

and representations from raw pixel data. CNN layers, such as convolutional, pooling, and fully connected layers, work together to detect and abstract features like edges, textures, and shapes, which are crucial for understanding image content. Feature fusion is a pivotal technique in leveraging the strengths of multiple feature sets derived from different models or sources. In this study, we employ CCA to fuse features extracted from two CNN models. This approach aims to enhance the representation of data by capturing the complementary information provided by each model. Once features are extracted, traditional machine learning algorithms, such as SVMs, random forest (RF), etc. can be employed for image classification.

This hybrid approach leverages the strengths of CNNs in feature extraction and the robustness of classical machine learning methods for classification tasks, often resulting in improved accuracy and performance in image analysis applications.

Deep CNN architecture. Classification is a crucial step in differentiating between disordered and healthy images. Recent studies have proposed various schemes for utilizing CNN models. One approach is to train the model using extensive datasets, while another involves using a pre-trained model through transfer learning. The proposed method employs a hybrid CNN model, comprising two main blocks: CNN-1 and CNN-2. These blocks first train on a large dataset of images, and then transfer the learned knowledge to subsequent blocks to assist in disease diagnosis. CNN-1 and CNN-2 contain 12 and 20 layers, respectively, including 2-dimensional Conv2D layers, and batch normalization layers. Additionally, they feature max pooling, dropout, and dense layers. Table 3 provides configuration details of both CNN models while Tables 4 and 5 provide a detailed overview of the layers for both CNN models.

In the proposed approach, we used CCA fusion to concatenate features extracted from both CNN models. The ensemble method's primary advantage lies in its ability to leverage the complementary strengths of individual models, potentially leading to improved diagnostic accuracy compared to using each model independently. The 20-layer CNN, for instance, includes batch normalization and additional layers that enhance feature extraction, while the 12-layer CNN may perform better in terms of computational efficiency.

Table 3. Configuration of both convolutional neural network (CNN) models.

Name	Parameter
Input	Fundus images from both dataset
Batch size	32
Optimization function	Adam optimizer
Image size	224 × 224 × 3
Loss function	Categorical cross-entropy
No of epochs	20
Activation function	Relu, Softmax
Dropout	40%

Feature extraction. Feature extraction is a critical component of the proposed model. In this study, we employed two CNNs to extract complementary features from retinal fundus images. CNN-1, structured with 12 layers, captures fundamental image patterns through a series of convolutional and pooling layers. Its output is a high-dimensional feature map, which is then flattened into a one-dimensional array for further processing. CNN-2, with its more complex architecture of 20 layers, including batch normalization and additional convolutional layers, is designed to capture more intricate features that are crucial for identifying subtle indicators of DR. Similar to CNN-1, the output of CNN-2 is also flattened, ensuring a consistent format for feature representation.

To effectively combine the features extracted from both CNNs, we employed CCA fusion. This method aims to identify linear combinations of the features from CNN-1 and CNN-2 that maximize their correlation, thereby enhancing the overall feature representation. The flattened feature vectors from both networks are input into the CCA algorithm, which computes canonical variables that best represent the combined features. The resulting fused feature vector retains the most relevant information from both models, providing a comprehensive representation for the final classification layer. This approach not only improves diagnostic accuracy but also leverages the unique strengths of each CNN in the detection of retinal diseases.

Table 4. An overview of 12-layer convolutional neural network (CNN).

Layer	Output Size	Parameters
Conv2D	(None, 222, 222, 32)	896
MaxPooling2D	(None, 111, 111, 32)	0
Conv2D-1	(None, 109, 109, 64)	18,496
MaxPooling2D-1	(None, 54, 54, 64)	0
Conv2D-2	(None, 52, 52, 128)	73,856
MaxPooling2D-2	(None, 26, 26, 128)	0
Conv2D-3	(None, 24, 24, 256)	295, 168
MaxPooling2D-3	(None, 12, 12, 256)	0
Flatten	(None, 36864)	0
Dense	(None, 128)	4,718,720
Dropout	(None, 128)	0
Dense-1	(None, 4)	516

Table 5. An overview of 20-layer convolutional neural network (CNN).

Layer	Output size	Parameters
Conv2D	(None, 222, 222, 32)	896
BatchNormalization	(None, 222, 222, 32)	128
MaxPooling2D	(None, 111, 111, 32)	0
Conv2D-1	(None, 109, 109, 32)	9248
MaxPooling2D-1	(None, 54, 54, 32)	0
Conv2D-2	(None, 52, 52, 64)	18,496
MaxPooling2D-2	(None, 26, 26, 64)	0
Conv2D-3	(None, 24, 24, 64)	36,928
BatchNormalization-1	(None, 24, 24, 64)	256
MaxPooling2D-3	(None, 12, 12, 64)	0
Conv2D-4	(None, 10, 10, 128)	73,856
MaxPooling2D-4	(None, 5, 5, 128)	0
Conv2D-5	(None, 3, 3, 128)	147,584
BatchNormalization-2	(None, 3, 3, 128)	512
MaxPooling2D-5	(None, 1, 1, 128)	0
Flatten	(None, 128)	0
Dense	(None, 512)	66,048
BatchNormalization-3	(None, 512)	2048
Dropout	(None, 512)	0
Dense-1	(None, 4)	2052

Let \mathbf{F}_1 and \mathbf{F}_2 denote the feature matrices extracted from these models

$$\mathbf{F}_1 \in \mathbb{R}^{N \times d_1} \quad (1)$$

where \mathbf{F}_1 is the feature matrix from CNN-1, where N is the number of samples and d_1 is the feature dimensionality.

$$\mathbf{F}_2 \in \mathbb{R}^{N \times d_2} \quad (2)$$

where \mathbf{F}_2 is the feature matrix from CNN-2, where d_2 is the feature dimensionality.

Feature selection through CCA fusion. The features extracted from the two CNNs are initially high-dimensional and contain

redundant information. To reduce redundancy and maximize relevant information, CCA is applied for feature selection. CCA works by identifying and preserving only those features that are maximally correlated across the two CNNs. This process selects a subset of features from each CNN output, ensuring that the final feature set includes only the most informative attributes of the retinal diseases. CCA fusion effectively captures disease-specific patterns by retaining features that show a high mutual correlation. Mathematically, CCA seeks to find transformation matrices W_1 and W_2 such that the transformed features Z_1 and Z_2 are maximally correlated.

$$Z_1 = F_1 W_1 \quad (3)$$

$$Z_2 = F_2 W_2 \quad (4)$$

CCA solves the following optimization problem

$$\max_{W_1, W_2} \text{corr}(F_1 W_1, F_2 W_2) \quad (5)$$

$$\text{s.t. } W_1^T F_1^T F_1 W_1 = I \quad (6)$$

$$W_2^T F_2^T F_2 W_2 = I \quad (7)$$

where corr denotes the correlation.

After determining the optimal transformation matrices W_1 and W_2 , we transform the original features.

$$Z_1 = F_1 W_1 \quad (8)$$

$$Z_2 = F_2 W_2 \quad (9)$$

The fused feature vector F_{fused} is obtained by concatenating Z_1 and Z_2

$$F_{\text{fused}} = [Z_1; Z_2] \quad (10)$$

where $F_{\text{fused}} \in \mathbb{R}^{N \times (d_1 + d_2)}$.

Classification with machine learning models. According to Baig,⁵⁴ the last step involves classifying testing images of fundus images to determine the kind of disease. The input image was defined in the proposed solution by selecting features and applying the multi-class classification approach. The objective in computing the categorization was to minimize computation time, so we employed machine learning algorithms. The classifiers we have utilized in this study include voting RF, COARSE-KNN, medium-KNN, fine-KNN, SVM, ensemble learning, gradient boosting machines, and Naive Bayes (NB). We utilized three machine learning algorithms as part of an ensemble classifier: SVM, logistic regression (LR), and decision tree (CT) classifier. Table 6 shows the hyperparameters for each model used for ensemble learning.

Results

We have conducted experiments to evaluate the proposed CNN model classification methodology, considering both qualitative and quantitative aspects. The evaluation involved testing the proposed method using the data we collected.

Table 6. Parameters for machine learning.

Model	Hyperparameter	Value
Support vector machine	Kernel	Radial basis function
	C	1.0
	Gamma	“scale”
Logistic regression	Solver	“lbfgs” (default)
	C	1.0
	Max_iter	100
Decision tree classifier	Criterion	“gini” (default)
	Splitter	“best” (default)

Findings on feature extraction utilizing CNNs

After training, the CNN can be used for feature extraction by feeding input data through the network and extracting the output of one of the intermediate layers. These extracted features can then be used as input to another machine-learning model or for further analysis and processing. This technique is often used in transfer learning, where a pre-trained CNN is fine-tuned on a new dataset for a specific task, leveraging the feature extraction capabilities learned from a large dataset.

In this section feature extraction results are given in both statistical as well as graphical form. In numerical form, accuracy, sensitivity, precision, recall, F1 score, and support are given using the formula given in equations (11) to (15).

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (11)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (12)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (13)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (14)$$

$$F1 - \text{Score} = 2 * \left(\frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (15)$$

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(FP + FN)(TN + FP)(TN + FN)}} \quad (16)$$

Accuracy, recall, and F1-score are crucial metrics for evaluating the performance of retinal image classification models.

While accuracy provides a general sense of overall performance, it can be misleading in imbalanced datasets, where high accuracy may result from simply identifying the majority class. Recall emphasizes the model’s ability to correctly identify positive cases, which is vital in medical contexts to ensure timely intervention for conditions that could lead to vision loss. The F1-score balances precision and recall, offering a comprehensive measure that helps mitigate both false negatives and false positives. Together, these metrics guide model selection, optimize decision thresholds, and facilitate continuous improvement, ultimately enhancing patient care and diagnostic accuracy in retinal disease detection.

In addition, the Matthews correlation coefficient (MCC) is also used for performance evaluation, as shown in equation (16). MCC estimates the correlation between predicted and actual value and is considered an important performance metric.⁵⁷

Feature extraction results using CNN-1. This section will delve into the results of feature extraction using CNN-1. The experiments employed the deep CNN base architecture model with training validation and testing data. Table 7 presents the statistical results of Feature Extraction from the CNN-1 model using data augmentation.

Figure 5 presents the accuracy and loss charts for CNN-1. It is observable in the charts that the model initiates with a starting training accuracy of zero, gradually advancing with increasing epochs.

Feature extraction results using CNN-2. This section will delve into the results of feature extraction using CNN-2. The experiments employed the deep CNN base architecture model with training validation and testing data. Table 8 presents the statistical results of Feature Extraction from the CNN-2 model using data augmentation.

Figure 6 presents the accuracy and loss charts for CNN-1. The plots show that the model starts with zero training accuracy and improves with time as the number of epochs increases.

Classification using machine learning model

The proposed methodology, which concatenates the features extracted by CNN-1 and CNN-2 into a single improved vector via CCA Fusion, was covered in earlier Sections. The classifiers were then given the fused vectors of the upgraded features to classify the input images. The proposed methodology includes several classifiers: Random forest, COARSE KNN, Medium KNN, Fine KNN, SVM, Ensemble Learning, Gradient boosting, and NB. In this step, several machine learning classifiers have been implemented. The goal was to reduce the overall system execution time as much as feasible.

Results for CCA fusion using random forest. A Random Forest classifier was used to categorize the abnormality from the fused feature vector. In this study, the accuracy of the DR class was 90.67%, the MH class was 91.18%, the ODC class was 94.5%, and the WNL class was 92.12%. In Table 9, different statistical parameters are cited. A 92.12% accuracy was achieved by this classifier.

Figure 7 illustrates the confusion matrix of RF showing 1520 correct predictions out of 1650 total predictions for four classes. In total 130 are wrong predictions by the RF model, with a higher number of wrong predictions for the DR class.

Results for CCA fusion using coarse-KNN. A coarse-KNN classifier was used to categorize the abnormality from the fused feature vector. In this study, the accuracy of the DR class was 88.28%, the MH class was 94.12%, the ODC class was 91.87%, and the WNL class was 86.7%. In Table 10, different statistical parameters are cited. A 90.24% accuracy was achieved by the Coarse-KNN classifier.

Table 7. Result for CNN-I network.

Diseases	Accuracy (%)
DR	87.09
MH	86.57
ODC	88.63
WNL	90.21

DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; CNN: convolutional neural network.

Figure 8 illustrates the confusion matrix for the Coarse-KNN. Results indicate that the performance of coarse KNN is better compared to RF concerning the MH class while poor for other classes. Overall, it correctly predicted 1489 instances while 161 instances were wrong showing its poor performance compared to the RF model.

Results for CCA fusion using medium-KNN. A Medium-KNN classifier was used to categorize the abnormality from the fused feature vector. In this study, the accuracy of the DR class was 87.32%, the MH class was 95.34%, the ODC class was 90.43%, and the WNL class was 86.95. In Table 11, different statistical parameters are cited. The model obtains a 90% accuracy using the CCA fusion.

Figure 9 illustrates the confusion matrix for Medium-KNN. Similar to coarse KNN, the performance of the Medium-KNN is better for the MH class with 389 correct predictions while poor for other classes compared to both Coarse-KNN and RF models. The model correctly predicts 1485 instances while 165 instances are wrong predictions.

Results for CCA fusion using fine-KNN. A Fine-KNN classifier was used to categorize the abnormality from the fused feature vector. In this study, the accuracy of the DR class was 86.6%, the MH class was 96.32%, the ODC class was 89.95%, and the WNL class was 86.95%. Table 12 shows various evaluation parameters for the Fine-KNN model. The model shows an accuracy of 89.94%. Compared to the RF model, it shows better performance for the MH class with a 96.32% accuracy while other classes have reduced values for accuracy, sensitivity, precision, and F1 score.

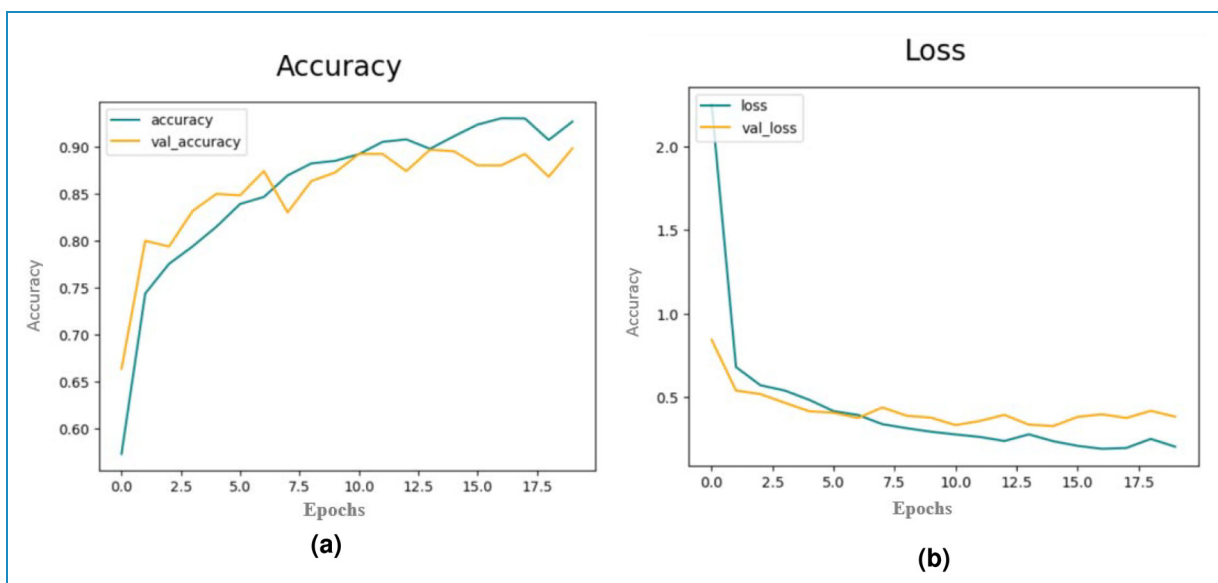


Figure 5. (a) Model accuracy graph for convolutional neural network (CNN)-I, and (b) Model loss graph for CNN-I.

Figure 10 illustrates the confusion matrix for Fine-KNN. Overall, 1484 instances are predicted correctly which is lower than RF, as well as, other variants of the KNN classifier. However, the number for correct predictions for the MH class is high which is 393. The model made 166 wrong predictions showing its poor performance compared to other models.

Table 8. Result for CNN-2 network.

Diseases	Accuracy (%)
DR	90.09
MH	91.56
ODC	89.53
WNL	91.21

DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; CNN: convolutional neural network.

Results for CCA fusion using SVM. An SVM classifier was used to categorize the abnormality from the fused feature vector. In this study, the accuracy of the DR class was 89.95%, the MH class was 92.4%, the ODC class was 94.74%, and the WNL class was 91.63%. In Table 13, different statical parameters are cited. 92.18% accuracy was achieved by this classifier.

Figure 11 illustrates the confusion matrix for SVM. With the SVM classifier, results for the MH, ODC, and WNL are better for all variants of KNN, as well as, the RF classifier with better accuracy and other parameters. The model made 1521 correct predictions, better than the previously used models' in this study while 129 predictions were wrong.

Results for CCA fusion using ensemble learning. An ensemble learning classifier was also used to categorize the abnormality from the fused feature vector. Using the ensemble classifier, the accuracy of the DR class was 92.34%, the MH class was 93.38%, the ODC class was 92.36%, and the

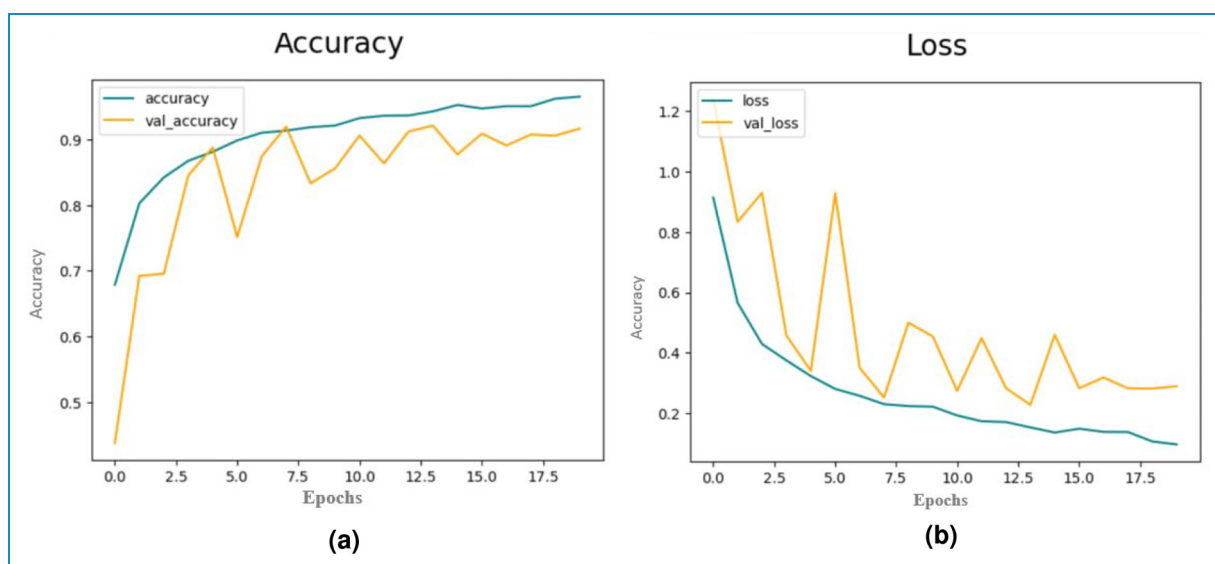


Figure 6. (a) Model accuracy graph for convolutional neural network (CNN)-2, and (b) Model loss graph for CNN-2.

Table 9. Class-wise statistics for CCA fusion using random forest.

Class	Accuracy	Sensitivity	Precision	Recall	FI Score	MCC
DR	90.67	90.67	94.75	90.67	92.66	90.14
MH	91.18	91.18	90.29	91.18	90.73	87.67
ODC	94.50	94.50	97.05	94.50	95.76	94.36
WNL	92.12	92.12	86.79	92.12	89.37	85.69

CCA: canonical correlation analysis; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; MCC: Matthews correlation coefficient.

WNL class was 92.12%. In Table 14, different statical parameters are cited. 93.39% accuracy was achieved by this classifier.

Figure 12 illustrates the confusion matrix for the ensemble classifier. The ensemble model performs superiorly with 1541 corrects predictions, better than all variants of the KNN, RF, and SVM classifier. The ensemble model makes only 109 wrong predictions for four classes.

Results for CCA fusion using gradient boosting. In addition to KNN, SVM, RF, and ensemble models, a gradient boosting classifier was also used to categorize the abnormality from the fused feature vector. In this study, the accuracy of the DR class was 92.12%, the MH class was 91.67%, the ODC class was 94.25%, and the WNL class obtained an accuracy of 92.36%. In Table 15, different statical parameters are cited. 92.36% accuracy was achieved by this classifier.

Figure 13 illustrates the confusion matrix for the gradient boosting classifier. With the correct prediction of 1528 instances, the model performs better than RF, Coarse-KNN, Medium-KNN, Fine-KNN, and SVM classifiers. The model has the second-best performance after the ensemble model with 122 wrong predictions.

Results for CCA fusion using NB. An NB classifier was used to categorize the abnormality from the fused feature vector. In this study, the accuracy of the DR class was 78.71%, the MH class was 84.07%, the ODC class was 85.17%, and the WNL class was 52.71%. In Table 16, different statical parameters are cited. 75.27% accuracy was achieved by this classifier.

Figure 14 illustrates the confusion matrix for the NB classifier. The performance of the NB classifier is very poor with only 1242 correct predictions and 408 wrong

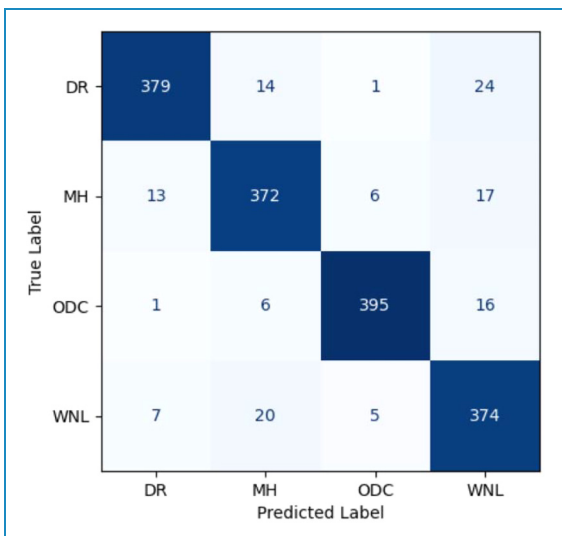


Figure 7. Confusion matrix for CCA fusion using RF. CCA: canonical correlation analysis; RF: random forest.

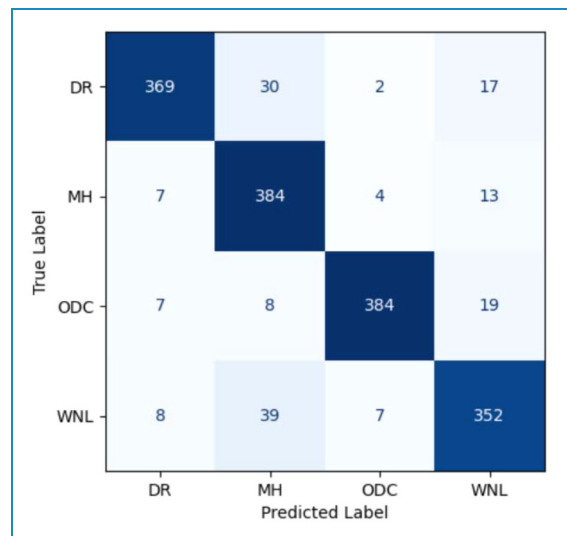


Figure 8. Confusion matrix for canonical correlation analysis (CCA) fusion using coarse-KNN.

Table 10. Class-wise statistics for CCA fusion using coarse-KNN.

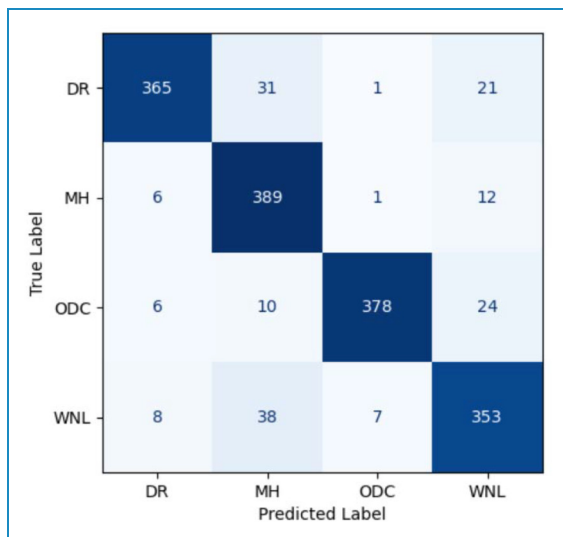
Class	Accuracy	Sensitivity	Precision	Recall	F1 Score	MCC
DR	88.28	88.28	94.37	88.28	91.22	88.46
MH	94.12	94.12	83.20	94.12	88.37	84.53
ODC	91.87	91.87	96.73	91.87	94.24	92.40
WNL	86.7	86.70	87.78	86.70	87.24	83.11

CCA: canonical correlation analysis; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; MCC: Matthews correlation coefficient.

Table 11. Class-wise statistics for CCA fusion using medium-KNN classifier.

Class	Accuracy	Sensitivity	Precision	Recall	FI Score	MCC
DR	87.32	87.32	94.81	87.32	90.91	85.31
MH	95.34	95.34	83.16	95.34	88.87	83.82
ODC	90.43	90.43	97.78	90.43	93.96	86.52
WNL	86.95	86.95	86.10	86.95	86.52	83.36

CCA: canonical correlation analysis; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; MCC: Matthews correlation coefficient.

**Figure 9.** Confusion matrix for canonical correlation analysis (CCA) fusion using medium-KNN.

predictions which are the highest among all models used in this study.

Performance comparison

This section compares the results of the suggested method to earlier research. We carried out pre-processing after obtaining the data, which included data augmentation to increase the size of the image collection. A more accurate model is trained as a result of this process, and overfitting is less likely to occur. The dataset contained images ranging in size from $512 \times 512 \times 3$ to $2144 \times 1424 \times 3$ to $4288 \times 2848 \times 3$. Each image has been resized to $224 \times 224 \times 3$. Crop the image to remove unwanted areas. Additionally, we utilized one-hot encoding during pre-processing to make categorical datasets easier to classify using CNN models by transforming them into numerical format. Using this encoding, categorical data can be

effectively learned from while keeping the relationships between them intact. Two CNN models are used to extract features from fundus images. Detailed information about the layers used in both CNN is provided in Tables 4 and 5. CNN-1 recorded a rather decent accuracy of about 89.52% for testing and approximately 89.85% for validation. The accuracy observed by CNN-2 for testing was around 91.39%, and for validation, it was about 91.67%. As discussed in the previous section, we trained two parallel CNN models to extract characteristics from the dataset. Each CNN produced recovered feature vectors, which we then turned on. Next, we used CCA fusion to extract and concatenate the most promising featured vector. Finally, for testing, we employ conventional machine learning models.

The classification results using the ensemble learning method reached 93.39%. For the computation complexity of the proposed approach, we kept it to a minimum by employing a conventional machine-learning model. While training traditional machine learning models only takes a few minutes, using a DL network at this stage could take several hours or even days. Using conventional machine learning techniques, we found that the ensemble learning model with CCA fusion was the most effective with an overall accuracy of 93.39%. Even so, out of all the classifiers we utilized in this study, NB provided the worst results. NB shows 75.27% accuracy using the CCA fusion. The performance comparison of all models is given in Table 17 without data augmentation. Table 18 shows statistical results for all models with data augmentation indicating better results when the models are trained using the augmented data.

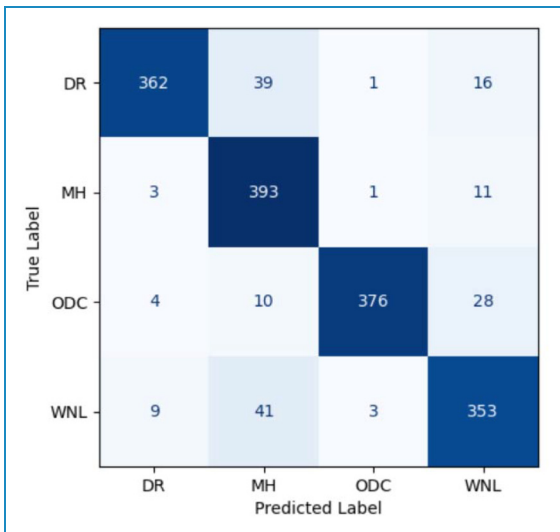
Performance concerning existing studies

The performance of the best-performing ensemble models is discussed concerning existing studies that utilized the same data. Table 19 shows the comparison of results obtained from the current study with exciting literature.

Table 12. Class-wise statistics for CCA fusion using fine-KNN classifier.

Class	Accuracy	Sensitivity	Precision	Recall	F1 Score	MCC
DR	86.60	86.60	95.77	86.60	90.95	88.82
MH	96.32	96.32	81.30	96.32	88.12	84.46
ODC	89.95	89.95	98.69	89.95	94.11	92.42
WNL	86.95	86.95	86.52	86.95	86.74	82.39

CCA: canonical correlation analysis; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; MCC: Matthews correlation coefficient.

**Figure 10.** Confusion matrix for canonical correlation analysis (CCA) fusion using fine-KNN.

The study¹¹ adopted a DL approach for fundus image classification using the RFMiD dataset and reported an accuracy of 89.10%. On the other hand,²¹ presented a CNN model and obtained a validation accuracy of 82.13%. The current study utilizes CNN models for feature extraction which are later fused to formulate a superior feature vector, thereby leading to an accuracy of 93.39% for RFMiD and RFMiD 2.0 datasets.

Discussion

This study provides an approach for better classification of fundus images using an ensemble model that combines SVM, DT, and LR classifiers. For model training, features obtained from two customized CNN models are used to make a single feature vector for better training of ML models. Experiments involve the original, as well as, augmented data. Extensive experimentation is carried out with a variety of ML models indicating the superior performance of the

proposed ensemble model. In summary, this study provides the following contributions.

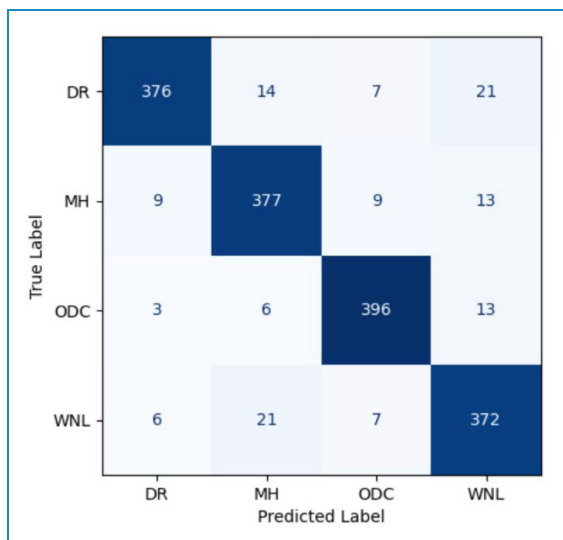
- **Improved Feature Representation:** The use of CCA for feature fusion effectively captured complementary information from the two CNNs, leading to a more robust feature set. This improved the overall classification performance compared to using individual CNNs.
- **Enhanced Classification Performance:** By employing machine learning classifiers on the fused features, the proposed model achieved higher accuracy and better generalization on the RFMiD and RFMiD 2.0 datasets. This demonstrates the potential of combining DL-based feature extraction with traditional machine learning classifiers.
- **Comparison with Existing Methods:** When compared to state-of-the-art methods in the literature, the proposed approach showed superior performance in terms of accuracy, precision, recall, and F1 score. This highlights the effectiveness of CCA fusion in enhancing the discriminatory power of CNN-derived features.
- **Versatility Across Datasets:** The model's performance was consistently strong across both RFMiD and RFMiD 2.0 datasets, suggesting that the proposed approach generalizes well across different variations of retinal fundus images.

In this study, we focused on feature extraction and fusion by employing two CNN models and combining their features using CCA fusion. The proposed methodology did not initially incorporate contrast enhancement or noise-handling techniques, as the goal was to assess the CNNs' raw feature extraction capabilities and the effectiveness of CCA fusion in retaining critical information for classification. The designed model enhances the model's ability to capture complementary information from fundus images. We implement an ensemble learning method that significantly improves classification accuracy compared to traditional methods, achieving an overall accuracy of 93.38%. The

Table 13. Class-wise statistics for CCA fusion using SVM classifier.

Class	Accuracy	Sensitivity	Precision	Recall	F1 Score	MCC
DR	89.95	89.95	95.43	89.95	92.61	90.27
MH	92.40	92.40	90.22	92.40	91.3	88.39
ODC	94.74	94.74	94.51	94.74	94.62	92.80
WNL	91.63	91.63	88.76	91.63	90.17	86.93

SVM: support vector machine; CCA: canonical correlation analysis; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; MCC: Matthews correlation coefficient.

**Figure 11.** Confusion matrix for CCA fusion using SVM. CCA: canonical correlation analysis; SVM: support vector machine.**Table 14.** Class-wise statistics for CCA fusion using ensemble classifier.

Class	Accuracy	Sensitivity	Precision	Recall	F1 Score	MCC
DR	92.34	92.34	94.37	92.34	93.79	91.14
MH	93.38	93.38	92.03	93.38	91.33	90.28
ODC	95.45	95.45	97.55	95.45	95.40	95.33
WNL	92.36	92.36	89.69	92.36	89.92	88.05

CCA: canonical correlation analysis; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; MCC: Matthews correlation coefficient.

proposed model demonstrates superior performance in detecting severe DR cases, with a precision of 94.37%, recall of 92.34%, and an F1 score of 93.79%. The proposed

approach addresses class imbalance through advanced data augmentation techniques, ensuring robust classification across different severity levels. This method is inspired by successful applications in other domains, such as lung cancer detection, where CNNs are used to extract spatial features from medical images and ML classifiers improve classification performance.³⁷ Additionally, research in medical imaging has demonstrated the effectiveness of using DL for feature extraction and combining it with traditional ML for classification, yielding more interpretable and computationally efficient models.³⁷ These contributions provide a scalable and accurate solution for automated DR localization and grading, facilitating early intervention and improving clinical outcomes.

Conclusion and future direction

The categorization of ocular illness is helpful in determining the eye's present state of health, analyzing the results of treatment, and choosing the best course of action. Creating a completely automated system is essential to enabling early identification and screening of people with eye diseases. A system of that kind ought to be non-invasive, reproducible, clinically dependable, and have a controllable decision-making process. Medical imaging and DL techniques present a viable way to provide comprehensive descriptions of diseases that have been identified. While we recognize the importance of computational complexity analysis, our focus in this manuscript is primarily on demonstrating the effectiveness of our proposed methodology in terms of classification accuracy, feature extraction, and overall model performance. However, several aspects of the proposed approach naturally contribute to managing computational complexity. First, during the preprocessing stage, we reduce computational overhead by resizing and cropping the images, which standardizes the input data and removes unnecessary information, thereby lowering the overall computational burden. Furthermore, we implement two CNN models, which are well-established for their ability to perform feature extraction efficiently while balancing accuracy and complexity. Additionally, by utilizing

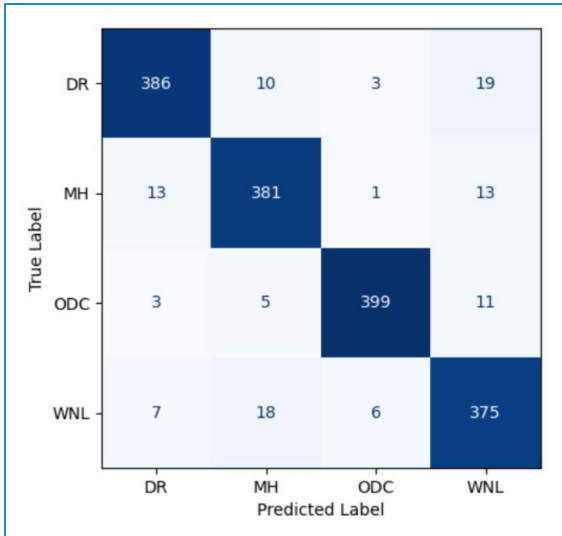


Figure 12. Confusion matrix for canonical correlation analysis (CCA) fusion using ensemble learning.

Table 15. Class-wise statistics for CCA fusion using gradient boosting.

Class	Accuracy	Sensitivity	Precision	Recall	F1 Score	MCC
DR	92.12	92.11	95.53	92.11	93.79	91.76
MH	91.67	91.67	91.00	91.67	91.33	88.47
ODC	94.26	94.26	96.57	94.26	95.40	93.88
WNL	92.36	92.36	87.62	92.36	89.92	86.58

CCA: canonical correlation analysis; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; MCC: Matthews correlation coefficient.

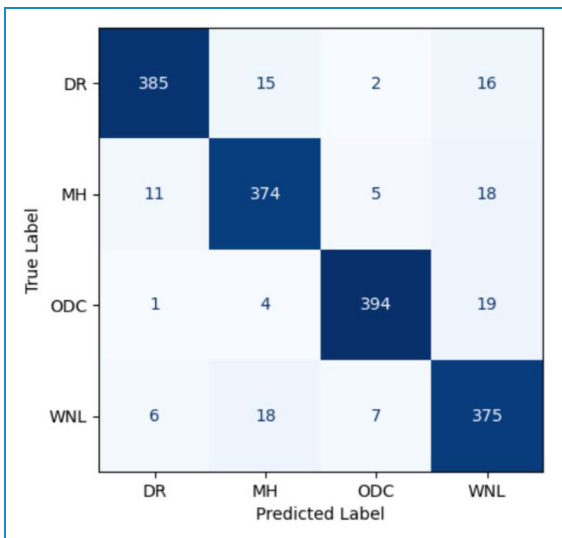


Figure 13. Confusion matrix for canonical correlation analysis (CCA) fusion using the GB classifier.

Table 16. Class-wise statistics for CCA fusion using NB classifier.

Class	Accuracy	Sensitivity	Precision	Recall	F1 Score	MCC
DR	78.71	78.71	85.17	78.71	81.83	76.26
MH	84.07	84.07	84.07	84.07	84.07	60.98
ODC	85.17	85.17	52.71	85.17	65.06	75.72
WNL	52.71	52.71	78.71	52.71	63.12	58.31

CCA: canonical correlation analysis; NB: Naive Bayes; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; MCC: Matthews correlation coefficient.

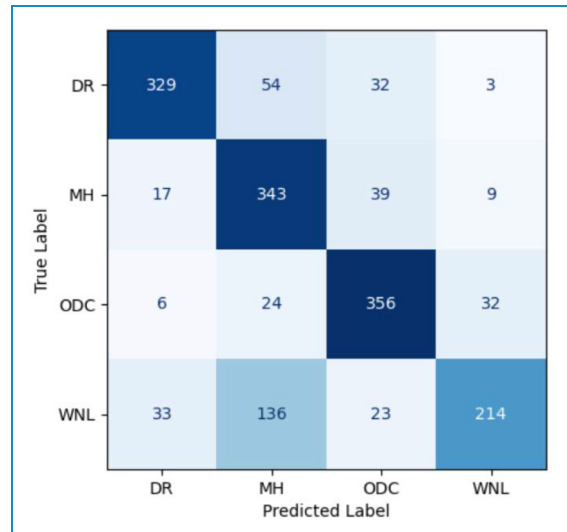


Figure 14. Confusion matrix for CCA fusion using NB classifier. CCA: canonical correlation analysis; NB: Naive Bayes.

CCA fusion, we avoid concatenating raw features, which could significantly increase the dimensionality of the feature space. Instead, CCA fusion optimally combines features from both CNNs, reducing redundancy and contributing to a more compact and manageable feature set, which in turn reduces the computational load during the classification phase. We also leverage traditional machine learning algorithms for classification, many of which are less computationally intensive compared to fully DL-based methods. This hybrid approach strikes a balance between complexity and performance, allowing us to benefit from DL’s feature extraction capabilities without imposing unnecessary computational costs during classification.

To assist in the diagnosis of various eye disorders, deep neural networks can develop hierarchical representations of images. However, due to the comparable look of fundus images of diverse diseases, diagnosing several eye ailments with a single neural network is difficult. Two CNNs are

Table 17. Class-wise statistics for all models using the original data.

Model	Accuracy	Classes	Acc. (%)	Sensit. (%)	Prec. (%)	Recall (%)	F1 (%)
RF	77.22%	DR	75.00	75.09	82.76	75.00	78.67
		MH	67.65	67.65	75.41	67.65	71.33
		ODC	30.00	30.00	52.94	30.00	38.46
		WNL	89.36	89.36	77.42	89.36	82.98
		Average	65.50	65.50	72.13	65.50	67.86
Coarse-KNN	71.72%	DR	67.71	67.71	70.65	67.71	69.15
		MH	69.12	69.12	61.84	69.12	65.29
		ODC	26.67	26.67	44.44	26.67	33.33
		WNL	81.91	81.91	78.57	81.91	80.19
		Average	61.35	61.35	63.88	61.35	61.99
Medium-KNN	70.94%	DR	67.71	67.71	77.38	67.71	72.16
		MH	67.65	67.65	57.50	67.65	62.17
		ODC	13.33	13.33	40.00	13.33	20.00
		WNL	82.98	82.98	75.00	82.98	78.82
		Average	57.92	57.92	62.47	57.92	58.29
Fine-KNN	70.94%	DR	62.5	62.5	85.71	62.5	72.46
		MH	73.53	73.53	55.56	73.53	63.32
		ODC	10.00	10.00	50.00	10.00	16.67
		WNL	84.04	84.04	73.15	84.04	78.28
		Average	57.52	57.52	66.11	57.52	57.68
SVM	73.82%	DR	70.83	70.83	78.16	70.83	74.33
		MH	61.76	61.76	67.74	61.76	64.64
		ODC	23.33	23.33	31.82	23.33	26.87
		WNL	87.77	87.77	78.1	87.77	82.65
		Average	60.92	60.92	63.95	60.92	62.12
GB	74.35 %	DR	79.17	79.17	73.08	79.17	76.02
		MH	57.35	57.35	68.42	57.35	62.3
		ODC	36.67	36.67	40.74	36.67	38.57
		WNL	84.04	84.04	81.44	84.04	82.72

(continued)

Table 17. Continued.

Model	Accuracy	Classes	Acc. (%)	Sensit. (%)	Prec. (%)	Recall (%)	F1 (%)
		Average	64.31	64.32	65.92	64.31	64.90
NB	59.69%	DR	41.67	41.67	74.07	41.67	53.19
		MH	29.41	29.41	50.00	29.41	37.04
		ODC	53.33	53.33	25.39	53.33	34.57
		WNL	80.85	80.85	67.56	80.85	73.68
		Average	51.32	51.32	54.25	51.31	49.62
Ensemble	75.92%	DR	76.04	76.04	75.26	76.04	75.65
		MH	67.65	67.65	69.69	67.65	68.65
		ODC	26.67	26.67	44.44	26.67	33.33
		WNL	86.70	86.70	81.10	86.70	83.81
		Average	64.27	64.27	67.62	64.26	65.36

SVM: support vector machine; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping

Table 18. Class-wise statistics for all models for augmented data.

Model	Accuracy	Classes	Acc. (%)	Sensit. (%)	Prec. (%)	Recall (%)	F1 (%)
RF	92.12%	DR	90.67	90.67	94.75	90.67	92.66
		MH	91.18	91.18	90.29	91.18	90.73
		ODC	94.50	94.50	97.05	94.50	95.76
		WNL	92.12	92.12	86.79	92.12	89.37
		Average	92.12	92.12	92.25	92.12	92.13
Coarse-KNN	90.24%	DR	88.28	88.28	94.37	88.28	91.22
		MH	94.12	94.12	83.2	94.12	88.37
		ODC	91.87	91.87	96.73	91.87	94.24
		WNL	86.70	86.70	87.78	86.70	87.24
		Average	90.24	90.24	90.52	90.24	90.26
Medium-KNN	90%	DR	87.32	87.32	94.81	87.32	90.91
		MH	95.34	95.34	83.16	95.34	88.87

(continued)

Table 18. Continued.

Model	Accuracy	Classes	Acc. (%)	Sensit. (%)	Prec. (%)	Recall (%)	F1 (%)
		ODC	90.43	90.43	97.78	90.43	93.96
		WNL	86.95	86.95	86.10	86.95	86.52
		Average	90.01	90.01	90.46	90.01	90.06
Fine-KNN	89.94%	DR	86.6	86.6	95.77	86.6	90.95
		MH	96.32	96.32	81.30	96.32	88.12
		ODC	89.95	89.95	98.69	89.95	94.11
		WNL	86.95	86.95	86.52	86.95	86.74
		Average	89.955	89.955	90.57	89.95	89.98
SVM	92.12%	DR	89.95	89.95	95.43	89.95	92.61
		MH	92.40	92.40	90.22	92.40	91.3
		ODC	94.74	94.74	94.51	94.74	94.62
		WNL	91.63	91.63	88.76	91.63	90.17
		Average	92.18	92.18	92.23	92.18	92.17
GB	92.61%	DR	92.12	92.11	95.53	92.11	93.79
		MH	91.67	91.67	91.00	91.67	91.33
		ODC	94.26	94.26	96.57	94.26	95.40
		WNL	92.36	92.36	87.62	92.36	89.92
		Average	92.60	92.6	92.68	92.6	92.61
NB	75.25%	DR	78.71	78.71	85.17	78.71	81.83
		MH	84.07	84.07	84.07	84.07	84.07
		ODC	85.17	85.17	52.71	85.17	65.06
		WNL	52.71	52.71	78.71	52.71	63.12
		Average	75.16	75.17	75.16	75.16	73.52
Ensemble	93.39%	DR	92.34	92.34	94.37	92.34	93.79
		MH	93.38	93.38	92.03	93.38	91.33
		ODC	95.45	95.45	97.55	95.45	95.40
		WNL	92.36	92.36	89.69	92.36	89.92
		Average	93.38	93.38	93.41	93.38	92.61

DL: deep learning; CNN: convolutional neural network; SVM: support vector machine; DR: diabetic retinopathy; MH: media haze; ODC: optic disc cupping; RF: random forest.

Table 19. Performance comparison with related existing studies.

Reference	Dataset	Model	Classes	Results
Das et al. ¹¹	RFMiD	Deep Learning	5	Accuracy 89.10 %
Kermary et al. ²⁵	RFMiD	Convolutional neural network (CNN)	4	Accuracy 76.04% Precision 76.81% Recall 76.14% F1-score 76.08%
Pandey et al. ²⁶	RFMiD,RFMiD 2.0	CNN	4	Accuracy 88.72% Precision 86.95% Recall 86.09% F1 score 86.13%
Proposed	RFMiD, RFMiD 2.0	Hybrid CNN and ensemble learning	4	Accuracy 93.38% Precision 93.41% Recall 93.38% , F1 Score 92.61%

used to extract the features. These derived features are concatenated using the CCA fusion method. The fused feature vector from the CNN models is used to train machine-learning models for better performance. By employing the ensemble learning classifier for fundus images, a 93.39% accuracy is obtained which is better than existing approaches. In the future, different fusion methods like serial fusion and principal component analysis can be applied to concatenate the features. We recognize the potential value of exploring contrast enhancement techniques, such as contrast-limited adaptive Histogram Equalization (CLAHE) or histogram equalization, to further improve the model's performance on low-contrast images. These techniques could be investigated as a preprocessing step to test their influence on feature extraction, particularly under suboptimal imaging conditions.

ORCID iD

Imran Ashraf  <https://orcid.org/0000-0002-8271-6496>

Statements and declarations

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is funded by the European University of Atlantic.

Conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability

The datasets utilized in this study, RFMiD and RFMiD 2.0, are publicly available and do not contain any personally identifiable information. Therefore, as per institutional and data-sharing guidelines, this study does not require ethics approval or waiver from an Institutional Ethics Committee (IEC) or Institutional Review Board (IRB).

<https://www.kaggle.com/datasets/andrewmvd/retinal-disease-classification>

References

1. Organization WH. Blindness and visual impairment. <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment> 2023.
2. Singh LK, Garg H, et al. Detection of glaucoma in retinal images based on multiobjective approach. *Int J Appl Evol Comput (IJAEC)* 2020; 11: 15–27.
3. National Eye Institute. Diabetic retinopathy. <https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/diabetic-retinopathy> 2023.
4. Xu K, Feng D and Mi H. Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image. *Molecules* 2017; 22: 2054.
5. Sengar N, Joshi RC and Dutta MK. An efficient artificial intelligence-based approach for diagnosis of media haze disease. In: *2021 12th International conference on computing communication and networking technologies (ICCCNT)*, pp.1–6. DOI: 10.1109/ICCCNT51525.2021.9579546.
6. Singh LK, Khanna M, Thawkar S, et al. Nature-inspired computing and machine learning based classification approach for glaucoma in retinal fundus images. *Multimed Tools Appl* 2023; 82: 42851–42899.
7. Singh LK, Garg H, et al. Detection of glaucoma in retinal fundus images using fast fuzzy C means clustering approach. In: *2019 International conference on computing, communication, and intelligent systems (ICCCIS)*, pp.397–403. IEEE.
8. Singh LK, Garg H, Khanna M, et al. An IOT based predictive modeling for glaucoma detection in optical coherence tomography images using hybrid genetic algorithm. *Multimed Tools Appl* 2022; 81: 37203–37242.
9. My Vision Organization. Optic nerve cupping. <https://myvision.org/eye-health/optic-nerve-cupping/> 2022.
10. Kar MK, Neog DR and Nath MK. Retinal vessel segmentation using multi-scale residual convolutional neural network (MSR-NET) combined with generative adversarial networks. *Circuits Syst Signal Process* 2023; 42: 1206–1235.
11. Das S, Lasker A, Ghosh M, et al. A deep learning-based approach for detecting diabetic retinopathy in retina images. In: *Internet of things-based machine learning in healthcare*. Chapman and Hall/CRC, pp.85–95.

12. Kaggle. Diabetic retinopathy detection. <https://www.kaggle.com/c/diabetic-retinopathy-detection> 2015.
13. Abramoff MD, Folk JC, Han DP, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmol* 2013; 131: 351–357.
14. Porwal P, Pachade S, Kamble R, et al. Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research. *Data* 2018; 3: 25.
15. Pachade S, Porwal P, Thulkar D, et al. Retinal fundus multi-disease image dataset (RFMiD): A dataset for multi-disease detection research. *Data* 2021; 6: 14.
16. Nguyen TD, Le DT, Bum J, et al. Retinal disease diagnosis using deep learning on ultra-wide-field fundus images. *Diagnostics* 2024; 14: 105.
17. Nagamani GM and Rayachoti E. Deep learning network (DL-NET) based classification and segmentation of multi-class retinal diseases using OCT scans. *Biomed Signal Process Control* 2024; 88: 105619.
18. Al-Fahdawi S, Al-Waisy AS, Zeebaree DQ, et al. Fundus-DeepNet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images. *Inf Fusion* 2024; 102: 102059.
19. Kumar KS and Singh NP. Retinal disease prediction through blood vessel segmentation and classification using ensemble-based deep learning approaches. *Neural Comput Appl* 2023; 35: 12495–12511.
20. Pustokhin DA, Pustokhina IV, Dinh PN, et al. An effective deep residual network based class attention layer with bidirectional LSTM for diagnosis and classification of COVID-19. *J Appl Stat* 2023; 50: 477–494.
21. Sengar N, Joshi RC, Dutta MK, et al. EyeDeep-Net: A multi-class diagnosis of retinal diseases using deep neural network. *Neural Comput Appl* 2023; 35: 1–21.
22. Pan Y, Liu J, Cai Y, et al. Fundus image classification using inception V3 and ResNet-50 for the early diagnostics of fundus diseases. *Front Physiol* 2023; 14: 160.
23. Choudhary A, Ahlawat S, Urooj S, et al. A deep learning-based framework for retinal disease classification. In: *Healthcare*, vol. 11, MDPI, p.212.
24. OpenGenus IQ. Understanding the VGG19 architecture. <https://iq.opengenus.org/vgg19-architecture/> 2020.
25. Kermany D, Zhang K and Goldbaum M. Large dataset of labeled optical coherence tomography (OCT) and chest X-ray images. In: *Mendeley data*. 2018, DOI: 10.17632/rsbjbr9sj.3. Version 3.
26. Pandey PU, Ballios BG, Christakis PG, et al. An ensemble of deep convolutional neural networks is more accurate and reliable than board-certified ophthalmologists at detecting multiple diseases in retinal fundus photographs. *Br J Ophthalmol* 2023; 108: 417–423.
27. Nguyen H. DiaretDB1: Standard diabetic retinopathy database. <https://www.kaggle.com/datasets/nguyenhung1903/diaretdb1-standard-diabetic-retinopathy-database> 2020.
28. Abhinav. Drishti-GS: Retinal image dataset for glaucoma screening. <https://www.kaggle.com/datasets/abhinav8617/drishti-gs> 2023.
29. Andrew Mvd. Drive: Digital retinal images for vessel extraction dataset. <https://www.kaggle.com/datasets/andrewmvd/drive-digital-retinal-images-for-vessel-extraction> 2019.
30. Budai A, Bock R, Maier A, et al. Robust vessel segmentation in fundus images. *Int J Biomed Imaging* 2013; 2013: 154860.
31. Linchundan. Fundusimage1000: Retinal fundus image dataset. <https://www.kaggle.com/datasets/linchundan/fundusimage1000> (2019, accessed 1 May 2024).
32. Andrew Mvd. Odir-5k: Ocular disease recognition dataset. <https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k> 2020.
33. Decencière E, Zhang X, Cazuguel G, et al. Feedback on a publicly distributed image database: The messidor database. *Image Anal Stereol* 2014; 2014: 231–234.
34. Zhang Z, Yin FS, Liu J, et al. Origa-light: An online retinal fundus image database for glaucoma analysis and research. In: *2010 Annual international conference of the IEEE engineering in medicine and biology*, pp.3065–3068.
35. Fu H, Li F, Orlando JJ, et al. Refuge: retinal fundus glaucoma challenge. 2019, DOI: 10.21227/tz6e-r977. <https://dx.doi.org/10.21227/tz6e-r977>.
36. Kaggle. STARE Dataset. <https://www.kaggle.com/datasets/vidheeshnacode/stare-dataset> 2019.
37. Thanki R. A deep neural network and machine learning approach for retinal fundus image classification. *Healthc Anal* 2023; 3: 100140.
38. Almustafa KM, Sharma AK and Bhardwaj S. Starc: deep learning algorithms' modelling for structured analysis of retina classification. *Biomed Signal Process Control* 2023; 80: 104357.
39. Ho E, Wang E, Youn S, et al. Deep ensemble learning for retinal image classification. *Transl Vis Sci Technol* 2022; 11: 39–39.
40. Ratul TR. ACRIMA Dataset. <https://www.kaggle.com/datasets/toaharahmanratul/acrima-dataset> 2023.
41. Elangovan P and Nath MK. Glaucoma assessment from color fundus images using convolutional neural network. *Int J Imaging Syst Technol* 2021; 31: 955–971.
42. Rodríguez MA, AlMarzouqi H and Liatsis P. Multi-label retinal disease classification using transformers. *IEEE J Biomed Health Inform* 2022; 27: 2739–2750.
43. Singh LK, Khanna M and Thawkar S. A novel hybrid robust architecture for automatic screening of glaucoma using fundus photos, built on feature selection and machine learning-nature driven computing. *Expert Syst* 2022; 39: e13069.
44. Xu W, Yan Z, Chen N, et al. Development and application of an intelligent diagnosis system for retinal vein occlusion based on deep learning. *Dis Markers* 2022; 2022.
45. Sunkari S, Sangam A, Suchetha M, et al. A refined ResNet18 architecture with swish activation function for diabetic retinopathy classification. *Biomed Signal Process Control* 2024; 88: 105630.
46. Elangovan P and Nath MK. En-ConvNet: A novel approach for glaucoma detection from color fundus images using

- ensemble of deep convolutional neural networks. *Int J Imaging Syst Technol* 2022; 32: 2034–2048.
47. Abitbol E, Miere A, Excoffier JB, et al. Deep learning-based classification of retinal vascular diseases using ultra-widefield colour fundus photographs. *BMJ Open Ophthalmol* 2022; 7: e000924.
 48. Kumar ES and Bindu CS. MDCF: Multi-disease classification framework on fundus image using ensemble cnn models. *J Jilin Univ* 2021; 40: 35–45.
 49. Panchal S, Naik A, Kokare M, et al. Retinal fundus multi-disease image dataset (RFMiD) 2.0: a dataset of frequently and rarely identified diseases. *Data* 2023; 8: 29.
 50. Vijayalakshmi D and Nath MK. A systematic approach for enhancement of homogeneous background images using structural information. *Graph Models* 2023; 130: 101206.
 51. Vijayalakshmi D and Nath MK. A strategic approach towards contrast enhancement by two-dimensional histogram equalization based on total variational decomposition. *Multimed Tools Appl* 2023; 82: 19247–19274.
 52. Vijayalakshmi D and Nath MK. A novel multilevel framework based contrast enhancement for uniform and non-uniform background images using a suitable histogram equalization. *Digit Signal Process* 2022; 127: 103532.
 53. Saponara S and Elhanashi A. Impact of image resizing on deep learning detectors for training time and model performance. In: *International conference on applications in electronics pervading industry, environment and society*. Springer, pp.10–17.
 54. Baig R, Rehman A, Almuhaimeed A, et al. Detecting malignant leukemia cells using microscopic blood smear images: A deep learning approach. *Appl Sci* 2022; 12: 6317.
 55. Majeed F, Shafique U, Safran M, et al. Detection of drowsiness among drivers using novel deep convolutional neural network model. *Sensors* 2023; 23: 8741.
 56. Mishra BK, Thakker D, Mazumdar S, et al. A novel application of deep learning with image cropping: A smart city use case for flood monitoring. *J Reliab Intell Environ* 2020; 6: 51–61.
 57. Anbalagan T, Nath MK, Vijayalakshmi D, et al. Analysis of various techniques for ECG signal in healthcare, past, present, and future. *Biomed Eng Adv* 2023; 6: 100089.