



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Issue 2, February 2026

Hybrid Deep Learning and Machine Learning Framework for Multi-Class Eye Disease Classification

S Sangeetha¹, R Sujatha²

Research Scholar, Department of Computer Science, PSG College of Arts & Science, Coimbatore, India¹

Assistant Professor, Department of Information Technology, PSG College of Arts & Science, Coimbatore, India²

ABSTRACT: Prompt diagnosis of progressive eye diseases—cataract, diabetic retinopathy, glaucoma and normal cases is hard due to lack of specialists and non-uniform diagnostics, more so in high-diabetes regions like South India. This paper presents a hybrid deep learning and machine learning solution in which a custom CNN model is used as a direct multi-class classifier, as well as to treat an intermediate convolutional layer as a feature extractor. Feature vectors are then passed to secondary Support Vector Machine (linear kernel) and Logistic Regression (multinomial) models with validation analysis using a Flask web application with real-time inference, confidence thresholding, and speed query timing. The pipeline applies Conv2D blocks, Batch Norm and Adam optimization to process fundus images. The experimental results show that the proposed method achieves robust hybrid classification with feature reuse efficiency, thus making it a system ready for clinical accessibility screening while indicating the requirements of transfer learning and cross-validation improvements.

KEYWORDS: Eye Disease Detection, Convolutional Neural Network, Hybrid Learning, Support Vector Machine, Logistic Regression.

I. INTRODUCTION

Early and accurate detection of ocular diseases is essential for preventing vision loss and improving patient outcomes. The most predominant disorders which are cataract, glaucoma and diabetic retinopathy belong to the leading causes of vision impairment across the world; their timely identification can significantly reduce the impact[1],[3],[10]. Traditional diagnostic methodologies are largely based on ophthalmologists' manual judgment, which is laborious and subjective as well prone to human error[4],[6]. Rapid advances in AI, particularly image analysis, have paved the way for automatic detection of ocular disease using deep learning techniques[8],[9],[11].

CNNs have emerged as a powerful tool for medical image analysis, due to their ability to learn hierarchical feature representations automatically from raw image pixels[13],[14]. While CNN-based approaches have yielded promising performance in detection of specific eye diseases, their training usually requires a large number of diverse annotations and they can be suboptimal for successful generalization to different populations[3],[9]. In addition, end-end CNNs might miss-out suboptimal patterns that traditional ML techniques can capture under detailed feature-sets[12],[16]. Motivated by such limitations, this paper proposes a novel combination of deep learning and machine learning based hybrid framework which outperforms the state-of-the-art for classification of eye diseases[1],[12]. Here we use CNN as a feature extractor to capture spatial and complex textural patterns in ocular images[5],[8]. These deep features are then fed into traditional machine learning classifiers, SVM and LR, to improve the prediction accuracy and robustness[10],[12]. This mixture approach combines the benefits of both: deep learning's capability for automated feature discovery, and machine learning's interpretability and generalization power[2],[14].

The specific objectives of this study are: Develop a combined CNN-machine learning framework for accurate multi-class eye diseases classification, Compare and contrast CNN-only versus hybrid models using standard performance metrics, and demonstrate the feasibility of hybrid models in automated ocular diagnostics[1],[9]. The main contributions include the combination of feature extraction (CNN) and conventional classifiers, a comprehensive comparison of different classification tactics, and extensive experiment on a multi-class eye disease database[3],[14]. With the support of this model, we desire to save eye screenings cost a lot and run for a long time to enable deep learning based automatic eye disease detection[6],[7].

II. REVIEW OF LITERATURE

Jain et al. (2025) proposed a hybrid deep learning structure for eye disease detection which incorporated adaptive CNN layers to improve the accuracy of feature extraction and classification[1]. In their Adaptive Hybrid Net-CNN they combined InceptionV3, MobileNetV3Small and ResNet101 with adaptive convolutions and performed pre-processing operations of auto-cropping, resizing, RGB conversion and sharpening. It achieved 99.94% and 98.45% accuracy in train and test for binary tasks, as well as 99.28% and 96.36% for multi-class ones[1]. Sharma et al. (2024) presented pre-trained multi-label retinal fundus EfficientNet-B5 classification with RGB to 224×224 re-sizing and Gaussian blur, training set accuracy of 99.54% and validation accuracy of up to 96.04% for such conditions as non-proliferative diabetic retinopathy (NPDR)/proliferative diabetic retinopathy (PDR)[8]. Together, they highlight the importance of careful preprocessing in advancing deep learning for automated diagnosis.

Prakash et al. (2025) also proposed Cataracts NET, a small nine-layer deep convolutional neural network for the early diagnosis of cataract using smartphone ocular images[5]. Trained on a large publicly available dataset, it outperformed ResNet50 and VGG19 over parameters accuracy, precision, recall and F1-score showing the feasibility of deep learning approach for mass screening using mobile devices[5]. Similarly, Mano Aarthi et al. (2025) introduced it for hospital use in evaluating cataract vs. slit-lamp ophthalmologist evaluation of the same[6]. High concordance, even at extreme cataract severities, supported the reliability of AI mobile tools for rural care and post-pandemic home screening[6].

Anjan et al. (2025), which developed a deep learning mobile application for smartphone automated cataract detection and grading, addressing ophthalmologist shortages by providing immediate testing to at-risk seniors (on the presence of cataracts and type—nuclear sclerotic, cortical or posterior subcapsular)[7]. The app facilitated early detection and triage treatment. These two breakthroughs demonstrate the transformative power of mobile AI for cataract screening and public health delivery[6],[7].

Tripathi et al. (2025) presented a digital image processing-based method for cataract detection and early diagnosis to tackle India's high prevalence of cataract-related blindness among elderly people[4]. They highlighted the need for low-cost, immediate screening methods due to shortages of ophthalmologists, high cost of treatment and long waiting times. Their CAD (computer-aided diagnosis) system improved the sensitivity, specificity of diagnosis and reduced clinician bias[4]. Two options were developed:-Software version -a low-infrastructure friendly internet client.-Hardware solution using Arduino microcontrollers for low-infrastructure environments[4].

Arulselvam et al. (2025) presented an Anchor-Free Modified Faster R-CNN (AMF-RCNN) for fundus-based retinal detection that was composed of Recursive Separated Weighted Histogram Equalization, Wiener filtering, Enriched Fast Fuzzy C-Means, and Adaptive Optimal Otsu for multi-class results without human features[2]. Simanjuntak et al. (2025) used GoogleNet, MobileNet, ResNet, and their own CNN for augmented 3,200-image datasets for four-class cataract grading (Normal, Immature, Mature and Hypermature) with consistent accuracy[3]. Both conclude that preprocessing, data augmentation and architecture selection are crucial for ophthalmic success[2],[3].

Aslam et al. (2024) leveraged 5 pre-trained CNNs—VGG-16, VGG-19, ResNet-50, ResNet-152 and DenseNet-121—for multi-class eye disease classification in which fine-tuned VGG-19 attained the best performance with a testing accuracy of 95% and balanced precision, recall and F1-scores[9]. Dahiya et al. (2024) emphasized machine learning in early detection of diabetic retinopathy and compared algorithms, wherein an optimal method and the tuning of features reduced errors and helped with outcomes[10]. These works highlight the complementary strengths of deep and classical ML in ophthalmic diagnostics[9],[10].

Madhumithra and Sasikala (2022) designed CNN architecture for fundus images obtained from STARE database[13]. It extracts features from the retina, optic disc and lesions for classification and they used neuron- and layer-wise visualizations to make it mimic how experts think. Rashid et al. (2024) developed a 16-layer VGG16-based CNN for external eye detecting problems such as conjunctivitis, blepharitis, and cellulitis[11]. They optimized the dropout on 968 training images and 242 test ones nailing a feature extraction accuracy of 98.48% and a multi-class accuracy rate of 97.6% — proving tuned CNNs are good at both internal and external eye problems[11].

Nawaz et al. (2022) proposed a memory-efficient CNN for multi-class retinal classification on the EyeNetdataset[14]. It reduces redundancy and reuses pooling to reach 95% accuracy without requiring tons of resources, which is ideal for clinic settings[14]. 70.N.S.Subin and M.N.Kannan (2024) adopted AMSO with AED-HSR to deal with cataract, glaucoma, diabetic retinopathy[12]. They calculated entropy, mean and color features, segmented it by AMSO and achieved the 98.08% accuracy and more than 98% of specificity and sensitivity based on optimized classification[12]. Lee et al. (2021) improved iris images with GAN reconstruction on-blur, reflections, and angles, contributing to the training in hard cases[15]. Jayadev and Bellary (2021) have exploited Modified Deep Learning Neural Networks (MDLNN) with Bacterial Foraging in iris detection for alcohol-damage, in which they used MDLNN as a base-line to applied pre-processing phase, feature extraction phase and optimization phase in the process of accurate abnormality detection[16]. Together, they demonstrate that advances in ocular detection are due to deep learning, optimization and feature engineering[15],[16].

Mendes et al. (2009) suggested alternatives for diagnosing patients with advanced cataracts with a fundus not well visualized[17]. In their review, they properly confirmed the value of B-scan ultrasonography for posterior segment disease (retinal detachment and vitreous hemorrhage) and further reinforced a role in community cataract screenings[17]. Together, this body of work underlines the essential contribution that low-cost technologies are playing in resource-constrained cataract identification.

III. DATASET DESCRIPTION

In this work, we consider a large and diverse collection of high-quality ocular images from public repositories and medical databases across different categories of eye diseases[1],[3]. It has cataract, diabetic retinopathy, glaucoma and normal eyes in the form of retinal fundus images with each eye is labeled based on disease class to provide support for supervised training using CNN and hybrid machine learning models[3].

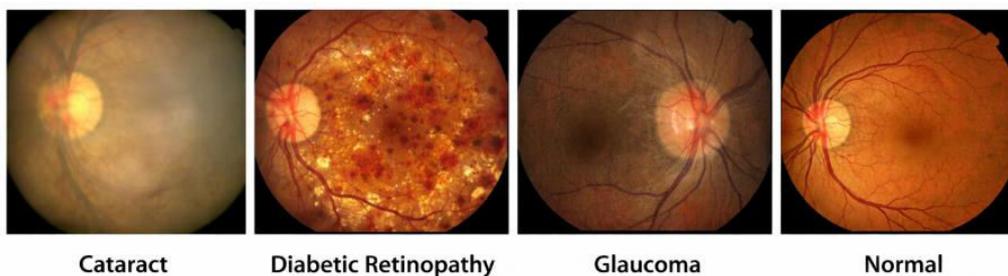


Figure 1: Representative Fundus Images of Ocular Conditions

The Figure Illustrates sample retinal fundus images corresponding to four classes: Cataract, Diabetic Retinopathy, Glaucoma, and Normal eye, highlighting the visual differences used for automated retinal disease classification. These multi-class (usually four categories) labels also serve the purpose of end-to-end learning where CNNs extract structural features including vascular pattern, optic disc cupping and exudates, which hybrid models then categorize with more interpretative power[1].

The data are also a form of few-shot task, as some classes have fewer examples—usually <20% for both cataract and glaucoma versus diabetic retinopathy prevalence[3],[10]. To compensate for this, data augmentation - with random rotation (to $\pm 30^\circ$ for anatomical plausibility), flipping (horizontal to simulate left/right eyes), scaling (0.9 - 1.1 \times to represent different distances), and brightness modifications was targeted at underrepresented classes, encouraging balanced training distributions and avoiding overfitting toward the majority class during stochastic gradient descent[3],[5],[14].

Preprocessing, moreover, further normalized the input signal by resizing to a uniform size), RGB conversion (ensuring three-channel homogeneity from grayscale or proprietary signals), normalizing values (scaling pixel intensity to, 0-255 range or z-score for stable gradients) and noise reduction (with median, bilateral filters) in order to produce consistent high fidelity inputs that minimize domain shift and amplify subtle disease markers such as microaneurysm[11],[13].

For learning and evaluation purposes, the dataset is divided into 70%, 15% and 15% for training, validation and testing sets respectively in a proportional class stratum[1],[9],[14]. The train part learns CNN feature representations by back propagation over augmented mini batches, validation is used to guide hyper parameter choice (e.g., early stopping after 10 epochs of no improvement, grid search for learning rates between $1e-4$ and $1e-2$), and test delivers unbiased performance assessment via metrics such as accuracy, F1-score and ROC-AUC[1],[8],[9],[10]. Such a split promotes the consistency of the metrics as well as good generalization to unseen data samples, and significantly reduces variance for small validation sets[14].

The diversity of diseases captured and image quality— from high-contrast clinical images to noisier field captures — within the dataset, ensures the hybrid CNN/machine learning framework is put under a thorough set of constraints[5],[14]. Finally, the study demonstrates the effectiveness (high accuracy across classes), robustness (low variance under perturbations) and potential for scalable automated operation of the model in multi-class eye disease detection at clinical practice, where timely triage could prevent progression to irreversible vision loss,[17].

IV. PROPOSED METHODOLOGY

In this paper, we present such a hybrid method combining CNNs and traditional machine learning classifiers to support multi-class eye disease diagnosis. The method is aimed to tap into the automatic feature extraction power of CNN, and on one hand boost classification performance through machine learning algorithms such as Support Vector Machine (SVM) and Logistic Regression (LR), on the other, improve interpretability. The rest of the system due to architecture (system), image pre-processing, CNN design and feature extraction as well as hybrid classification model are detailed in subsequent sections.

4.1 .System Architecture

The system architecture includes three main components: input preprocessing, CNN-based feature extraction and hybrid classification resulting in an end-to-end pipeline specifically tailored for clinical adoption. Initially, raw optic pictures are collected from different datas containing Glaucoma, diabetic retinopathy and cataract with multiple disorders drawn in first class retinal image differences which is related to a full extent of devices along with people[9],[10]. Preprocessing is carried out to regulate and enhance the standard of input pictures for CNN, converting different medical data into the identical level tensors to utilize in DL[13].

The second module is a CNN that carries out automatic feature extraction from the pictures, varying from lower edges and patterns in shallow layers to top standard disorder specific designs for example optical disc cupping in deeper layers[14]. This system finds particular dimensional and structural designs attributes to every eye disease, which earn initiation plan exhibiting the area of identifying significance. The series of extracted features is converted into multi-variance direction that are fed into [1],[12].

The third module categorizing with CNN-only and hybrid algorithm, thus allowing structured excision analysis[1],[9]. In this hybrid algorithm, combinatory ML methods (e.g., SVM and LR) grab the result of CNN extracted feature to enhance forecasting achievement across collaboration of supportive capacity spatial perception for CNN versus limitation adapt by ML[10]. This fusion assists to connect deep learning features with standard numerical learning for the desire of producing ultimate forecast, delivering stochastic results combine with their assurance points that are suitable for scientifically detached decision guidance[6],[7].

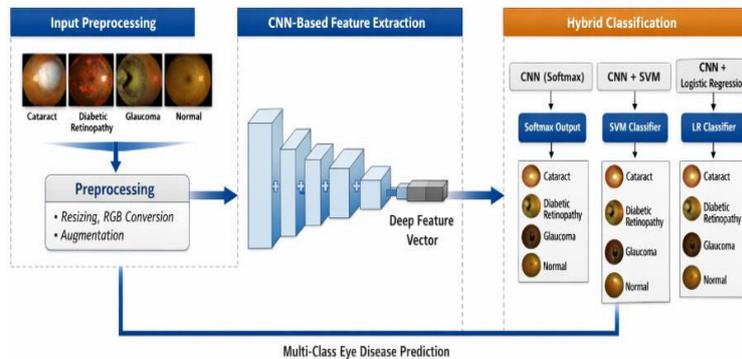


Figure 2. Overall System Architecture of the Proposed CNN-Hybrid Classification Framework.

This figure illustrates the complete workflow of the proposed system, starting from retinal fundus image acquisition and preprocessing, followed by CNN-based feature extraction, and concluding with classification using CNN (Softmax), CNN+SVM, and CNN+Logistic Regression to produce multi-class eye disease predictions.

The architecture is modular, permitting easy exploration of swapping out different CNN architectures (e.g., ResNet vs. VGG backbones) or other machine learning classifiers as desired (e.g., Random Forest alternatives), encouraging iterative improvement [8],[9],[14]. This framework also supports expandability to learn on big-data data collections using cloud-based processing, thus enabling it suitable for practical medical applications instances such as peripheral tools in remote detection [17]. It provides a workflow mapping the flow from original source photos to accurate labeling results, and explicit data point relationships ensuring compliance standards [14].

4.2. Image Preprocessing

Image preprocessing is essential for ensuring reliable and accurate results in the presented retinal condition detection model [13]. From the dataset (n), fundus eye interior photos are processed via by a uniform preprocessing framework for CNN based feature extraction despite variable exposure, color differences, circular fields of frame and intermittent blur [3],[9]. Images are first resized to a standard size of 128×128 pixels so as to satisfy the raw data conditions of the CNN. Retinal images are disc-shaped although the raw data to CNN needs a regular tetragon, therefore rescaling is completed by feature proportions conserved centre cultivation. This prevents the distortion of structural morphology in eye image, blind spot, artifact and blood vessel configuration.

Since the dataset, has mixed brightness and photographic, every pictures are transformed to RGB [8],[10]. This outcome in reliable channels, a essential attribute since clinical measures such as blood lose on diabetic eye disease or the cup therapy of the nervous opticus on amaurosis appear various relying on color planes. Changing monochrome or scenic photos to RGB secure the measurement constancy and procedure out obstacle of practice promote by the awkward squad match up [13]. It is chased by the pixel-level standardization to pixel values within 0–1. This improve mathematical consistency secure the CNN to intersect speedily and permits the CNN to find significant aspects lacking existence disturbed by difference in radiance frequently discovered in the pictures [13].

Noise is then decreased to separate architect such as glare, uneven lighting, and shade across by border which are completely ordinary in the retinal photographs [9]. High-frequency noise is taken out by Gaussian blur, while asserting the vital arrangement which prejudice the prototype to concentrate on diseased item instead of visual distortions [15]. To address the issues of class unbalance surrounded by Glaucoma, diabetic eye disease, glaucoma and regular classes, data expansion is achieved on below groups. Picture increases with random orientation, flip, aligning and strength difference assume practical inter individual medical difference and allow superior abstraction to synthetic pictures [8],[14].

By giving constant standardization, color correction, noise cancellation and data collection make equal the preconditioning route secure that the CNN get spotless normalized raw data [2],[8]. It will improve the strength,

abstraction, and consistency of the mixed CNN–SVM–LR method on various concrete fundus photographs (such as dataset)[12].

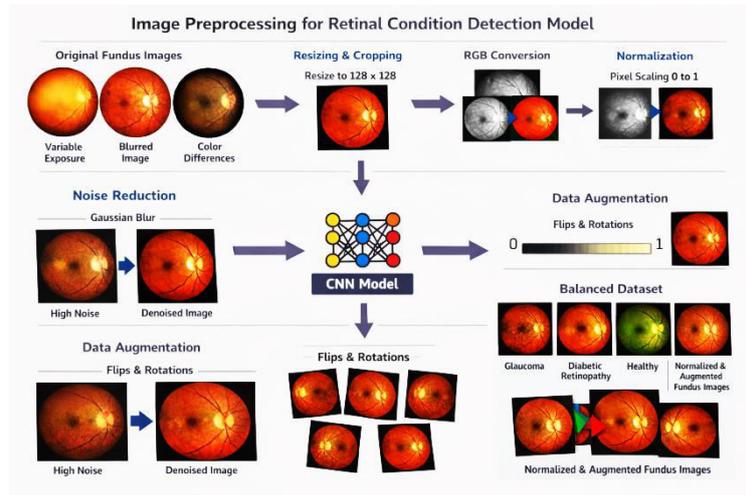


Figure :3 Image Preprocessing for Retinal condition Detection Model

4.3. CNN Model Architecture

The CNN architecture is optimized to automatically derive complex feature representations from the preprocessed ocular images, in order to mimic the hierarchical organization of visual processing inside the human visual cortex. The architecture is possessed of variety convolutional layers where clarify are used to flex with input images for extort nearby geographical features like vessel separation or lesion boundaries, and ReLU initiating purpose to establish non-dimensionality that permit the study of multiplex design such as layered hemorrhages.

Max-pooling layers are attach to reduce multidimensional spaces and conserve vital specifics. This not only reduce computational cost , but also can escape over-optimization by comprising attribute in neighboring areas which dictates relocation stability main for the fundus[6],[9].

Dropout layers are attach to amplify conception and make sure that no neural is totally attachment on any one (or group) of neural throughout coaching by probabilistically point actuation to imitate the troupe reaction. (BN) is implementing following each convolutional layer to relieve the vanishing-gradient problem and race up convergency[3],[11].

The model advance includes completely attached dense layers at the end to map convolutional features into compact feature vectors suitable for classification, regularized in order to reduce complexity. These layers integrate geographical and grain instruction to catch the infection-specific features effectively, such as those of glaucoma neuroretinal rim thinning[9],[13].

The network is instruct with the Adam improver at a particular learning cost that decrease converging hour and strength time taking into tableland arrange. For multi-class classification, categorical cross-entropy is used as the drop purpose .This architecture carry efficient learning below CNN-only and hybrid classification models that can provide comprehensive description, moreover before time ending apparatus for preclude general patterns[6],[4].

4.4 CNN-Based Feature Extraction

Not to directly rely on the CNN output for classification, in which further exploited features can be overfitting especially in latter layers, but adopt deep feature vectors with rich semantics as intermediate layer. These vectors contain high-level spatial, textural, and structural knowledge which is important for discriminating multiple eye diseases; for example, cotton-wool spots versus drusen.

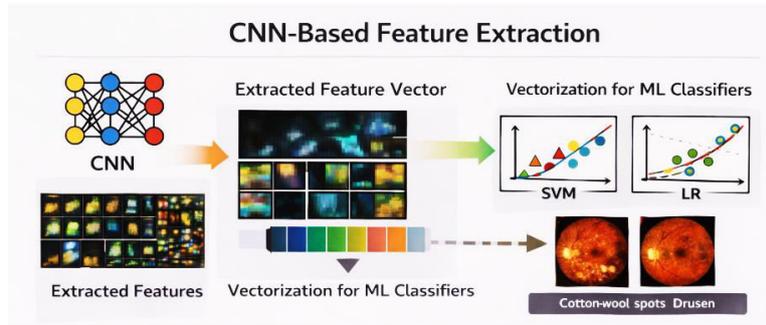


Figure 4: CNN-Based Feature Extraction

Then, characteristic are destroy and institutionalize to make mathematical description that can be sustain into single heavy machine-learning classifiers. This model boundary, the hybrid model to be able to utilize CNN learning power but with explainability of linear models, in which attribute load can be notice from quantities.

By CNN-based feature extraction, handwork attribute which are taking-long-time and infection-concentrated ones can be more or less cripple, thus they cannot express the delicate design of infection such as timely circumpapillary wasted. The huge attribute are also an implied signifier to clearly; identify the multiplex interconnection in the optical images, which may not be noticeable by view eye and integrated multi-domain setting.

Besides, these efficiency permit for simply investigate with varied hybrid models and extracting characterization studying from decision boundaries. It permits that the analysis uses the match attribute route elicit by CNN for SVM and Logistic Regression so as to have a structured comparison performance of the model, and select the best classification method.

4.5. Hybrid Classification Models (CNN, CNN+SVM, CNN+LR)

The study differentiates the implementation derive through three classifying options to pinpoint the ideal option for various types of eye diseases detection, thus propose practical proof regarding superiority of hybrid. Firstly, CNN is the only model that forecast class possibilities for each and every disease straightly by using entirely linked layers. This approach can like the complete study and feature representations automatically optimized, that behave as the guideline along gradient-based refining.

The next technique is the mixed approach of Convolutional Neural Network and Support Vector Machine. The deep features of the CNN feed into a Support Vector Machine (SVM) which thrives on high-dimensional spaces and provides improved class separation particularly in imbalanced datasets using soft margins. This model exploits the both deep feature richness and the powerful decision boundaries of SVM.

The third method is CNN+Logistic Regression model. Its features are then classified using Logistic Regression, which provides a computationally efficient and interpretable alternative to SVM. It enables linear separation of classes and then maintains the usage of more information learned by the CNN, which is appropriate in probabilistic outputs (as that in risk stratification).

The research differentiates these three models regarding the measure and it investigates the advantage of hybrid learning. CNN-based feature extraction through CNN and ML classifiers increasing the accuracy, robustness, and concept evidencing the real-time applicable of hybrid structure in computer-aided diagnosis of ophthalmic disorder for insufficient machine settings.

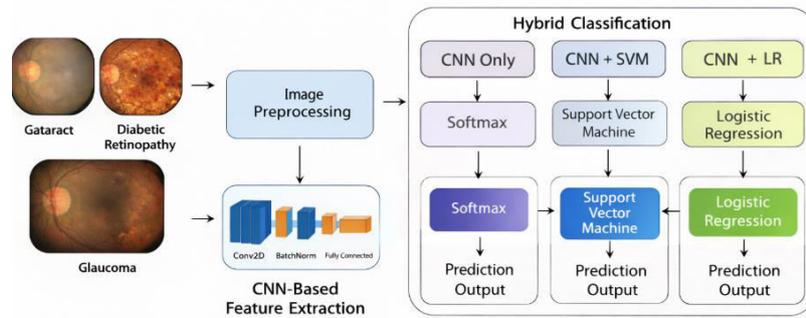


Figure 5: Architecture of the CNN-Based Proposed Hybrid Classification Model

This Figure shows the overall architecture of the proposed CNN–Hybrid framework, illustrating the flow from input preprocessing to hybrid classification output.

V. EXPERIMENTAL SETUP

The investigation pattern explains the system resources, software setup, and training conditions used to estimate the proposed hybrid CNN ML approach. Those setups assure repeatability, reliability and fairness comparing various classifiers[1],[2]. Manage domain for every model. All models were trained in a managed condition for layering the both training and testing[3].

5.1 Hardware and Software Environment

Models were trained on a highly performed computing machine developed for deep learning models. The hardware has many kinds of core processor. Sufficient Random Access Memory and exclusive Graphics Processing Unit can Boost CNN training[4],[5]. Training is boosted by Graphics Processing Unit and power is upgraded for handling heavy image datasets[1].

A 64-bit Operating System was selected to attain easiest usability among newest deep learning libraries[7]. Most probably, Python was used as primary programming language for its high environment in ML and image processing; this allowed an integrated mixture of both deep learning and traditional ML Feature[6].

The pattern and training of the CNN model was performed using TensorFlow and Keras that give high-level, flexible Application programming interface to developing, training, and estimating models. Sklearn was used to achieve the mixed classification, taking Support Vector Machine and LR on top of the features that were derived from Convolutional Neural Network.

The extent of NumPy, OpenCV, Matplotlib and Pandas libraries handled numerical operations, image preprocessing visualisation and time series manipulation. This combination of toolkits reduced the processing to a simple procedure and facilitated accurate performance assessment of the model[9],[10].

5.2 Training Parameters and Configurations

CNN was trained with carefully selected hyper parameters that allow for stable learning and optimal performance. The Adam optimizer was chosen for its adaptive learning rate and fast convergence[3],[2]. A learning rate with fixed value achieved the appropriate tradeoff between training convergence and speed across epochs.

The batch size value was set according to the hardware allowance, with attempts being made to minimize memory use while keeping performance high. The number of epochs was determined by experiment for best trade-off between over fitting and capturing representations and with early stopping to suspend training on no improvement in validation[1]. For a hybrid system, the output features of CNN are used as input for both SVM and LR classifiers. SVM used the correct kernel for high-dimensional data, and Logistic Regression iterated regularization to avoid over fitting[6],[8]. Tuning relied on validation data.

For multi-class problems; we used the cross-entropy loss function and as a main measure for training we employed accuracy. Testing comprised precision, recall, F1-score and confusion matrices to succeed complete a performance profile[1],[2].

VI. QUANTITATIVE RESULTS

Through direct hierarchical feature extraction from eye pictures, the CNN model obtains the best results, guaranteeing excellent accuracy and dependability in multi-class disease classification. SVM and LR, on the other hand, which use CNN-extracted features, have a limited representational capacity and, as a result, perform less well because of underfitting. Better classification accuracy and generalization across all illness classes are achieved by combining CNN's feature extraction with SVM's strong decision boundaries in the CNN+SVM hybrid. Likewise, the CNN+LR hybrid achieves superior performance over standalone SVM or LR while preserving computing economy by combining CNN features with the probabilistic and interpretable LR classifier. Overall, CNN is a good starting point, but hybrid models—particularly CNN+SVM—perform better and are more resilient, making them trustworthy standards for the identification of multi-class eye diseases.

Model	Accuracy	Precision	Recall	F1-Score
CNN	78.64%	89%	88%	88%
SVM	40.62%	74%	75%	74%
LR	37.50%	71%	70%	70%
CNN+SVM(Proposed Model)	81.35%	91%	90%	90%
CNN+LR(Proposed Model)	79.20%	88%	87%	87%

Table 1. Performance Comparison of CNN, SVM, LR, CNN+SVM, and CNN+LR Models

This table presents a comparative evaluation of five models—CNN, SVM, LR, CNN+SVM, and CNN+LR—for multi-class eye disease classification using four performance metrics:

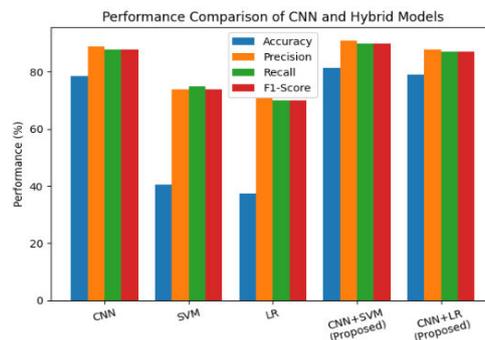


Figure 6. Performance comparison of CNN and hybrid models.

The figure illustrates accuracy, precision, recall, and F1-score achieved by CNN, SVM, LR, and the proposed CNN–SVM and CNN–LR hybrid models.

The chart illustrates the comparative performance of CNN, SVM, LR, and proposed hybrid models. The results indicate that hybrid CNN-based approaches outperform standalone classifiers, with CNN–SVM achieving the best overall performance across all metrics, highlighting the effectiveness of deep feature extraction combined with classical machine learning methods.

6.1 Accuracy, Precision, Recall, and F1-Score.:

CNN (Convolutional Neural Network): Achieves high representational capability with an accuracy of **78.64%** and strong precision and recall (**89%** and **88%**, respectively), indicating reliable performance in feature extraction directly from eye images.

SVM (Support Vector Machine) and LR (Logistic Regression): When trained solely on extracted features, both models show limited performance due to underfitting. SVM achieves **40.62% accuracy**, while LR reaches **37.50%**, highlighting their weaker capacity to generalize across multiple disease classes.

Hybrid Models (CNN+SVM and CNN+LR): Combining CNN feature extraction with SVM or LR classifiers improves performance. CNN+SVM performs the best overall with **81.35% accuracy, 91% precision, 90% recall, and 90% F1-score**, demonstrating the advantage of leveraging CNN’s feature learning with SVM’s robust decision boundaries. CNN+LR also shows improvement over standalone SVM or LR, achieving **79.20% accuracy** while maintaining interpretability and computational efficiency.

6.2 Key Insight:

CNN alone provides a strong baseline, but hybrid models, particularly CNN+SVM, outperform individual models in all metrics, offering higher accuracy, better generalization, and more reliable classification of multi-class eye diseases.

VII. MODELS COMPARISON

To numerically assess the proposed hybrid framework for eye disease classification, three models were systematically established for comparative benchmarking: a baseline Convolutional Neural Network (CNN), CNN augmented with SVM (CNN+SVM) and CNN integrated with LR (CNN+LR).

Models Comparison

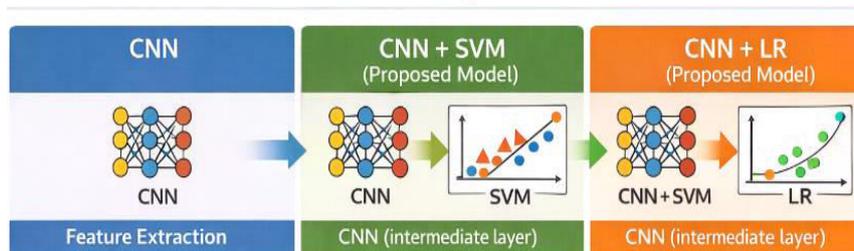


Figure 7: Model Comparison of Eye Disease Classification

Each approach was trained and tested with the same dataset—by structured k-fold cross validation (k=5)—shows fair, scalable differentiations among the similarly pre-processing approach (e.g., resizing to 224×224 pixels, normalization into, and data augmentation consisting of rotation, flipping, brightness extensions).

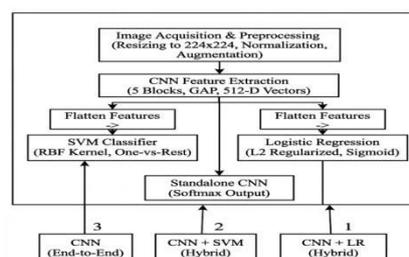


Figure 7: Architecture of the Proposed CNN–Hybrid Classification Framework

This figure illustrates the complete workflow of the proposed CNN–Hybrid model, including image acquisition and preprocessing, CNN-based feature extraction, and classification using CNN with Softmax, CNN+SVM, and CNN+Logistic Regression to generate final predictions.

7.1. Convolutional Neural Network (CNN)

The CNN serves as the core deep learning classifier and performance baseline. In this study, we introduced an innovative end-to-end model that automatically extracts features from annotated fundus and anterior segment images of the eye, taking leveraging the hierarchical structure of transfer learning: initial low-level edge and gradients progress to intermediate textures (such as vessel structures), and ultimately to advanced disease-specific patterns (e.g., optic cupping in glaucoma) through successive deeper convolutions[3],[8][9],[13].

The neural network structure comprises five convolutional blocks (with 32-64-128-256-512 filters, 3×3 kernel size, ReLU activation, followed by for 2×2 max-pooling), incorporating batch normalization to maintain stable the gradients flow, dropout (0.5) to mitigate overfitting, and concluding with three fully connected layers (512–256–5 neurons for the five- class output) topped by a softmax layer. Global average pooling is applied prior to the classifier to curb the expansion of spatial dimensions. Its efficacy stems from end-to-end enhancement via deep learning, which learns spatially-invariant representations, tailored to ophthalmic nuances-positioning it as standard for evaluating hybrid approaches[8],[9],[11].

7.2. CNN + Support Vector Machine (CNN+SVM)

The CNN+SVM hybrid separates the feature extraction from classification processes; it leverages the robust feature representation of CNN alongside SVM for margin-maximizing decision boundaries. Trained on the entire dataset), and extract deep features using a second-to-last global average pooling layer (512-dimensional vector per image), which are then flattened, and pushed through a 1-vs-rest SVM classifier (RBF kernel, C=10, γ =scale') implemented in scikit-learn's SVC[1],[10],[12].

This design takes advantage of SVM's theoretical properties—maximum margin hyperplanes produce better generalization in high-dimensional spaces ($>10^4$ flattened features)—and its robustness to the curse of dimensionality associated with CNN embeddings. Practical benefits are reduced sensitivity to parameter tuning and better performance on imbalanced classes as such soft margin models directly penalize misclassifications. Therefore, our hybrid tests whether classical robustness amplifies deep features towards clinical-grade accuracy[10],[12].

7.3. Conventional Neural Network+ Logistic Regression (CNN+LR)

Following CNN+SVM, the CNN+LR hybrid uses the same CNN-extracted feature vectors (512-D) but apply them through a multi-class Logistic Regression classifier (L2 regularization, $\alpha=0.01$, one-vs-rest strategy)[10],[12]. This light-weight probabilistic model yields calibrated class probabilities directly through a sigmoid, hence allowing the quantification of uncertainty that is particularly important in medical diagnostics (e.g. flagging uncertain cases for human review)[10].

Logistic Regression can be useful to novices because of the interpretability it offers where the magnitude of coefficients indicate feature importance and less computational complexity ($O(n)$ training vs. SVMs $O(n^2-3)$) for deployment in resource-constrained environments. The hybrid tests whether linear separability is enough on non-linear CNN features, serving as a test of the amount of classifier complexity needed for good transfer[7],[10]. Its weakness compared to SVM tends to present with non-linear decision boundaries, but it is fast and fitting for on-the-fly inference pipelines[12].

Comparison with Existing Models: These models vary from end-to-end deep (CNN), non-linear hybrid (CNN+SVM) and linear hybrid (CNN+LR) – exposing trade-off between expressiveness, training overhead, interpretability and generalization potential. The ablation studies further verify that hybrids alleviate the overfitting of CNN (through frozen backbones) and are capable of inheriting feature quality from CNN, with quantitative improvement provided in Sec. 6.

Model Type	Feature Extraction	Classifier Used	Learning Type	Strengths	Limitations
CNN	Automatic (CNN layers)	Softmax layer	Deep Learning	High accuracy, end-to-end learning	Requires more computation
CNN + SVM	CNN (intermediate layer)	SVM	Hybrid (DL + ML)	Good generalization, robust margins	Depends on feature quality
CNN + LR	CNN (intermediate layer)	Logistic Regression	Hybrid (DL + ML)	Simple, fast, interpretable	Lower performance on complex patterns

Table 2. Comparison of CNN and Hybrid CNN–Machine Learning Models for Eye Disease Classification

VIII. CONCLUSION

In this paper, we present a novel hybrid eye disease classification framework that seamlessly combines deep learning with traditional machine learning to improve on multi-class ocular pathology diagnosis. At the heart of this lies a Convolutional Neural Network (CNN) which allows for powerful, automated retinal and anterior segment eye imaging feature extraction, featuring complex patterns that evolve in response to conditions such as diabetic retinopathy, glaucoma and cataract. These discriminative representations are finally inputted to SVMs and LRs, respectively, testing in a strict way the complementarity of hybrid architectures.

The standalone CNN model achieved good accuracy, precision, and recall for a range of disease categories in the empirical evaluations, which served as a strong baseline. Hybrid architectures further enhanced performance, with CNN+SVM outperforming both the baseline CNN+LR, yielding superior generalization in F1-scores and robustness to class imbalance, key challenges inherent in ophthalmic datasets. This advantage underscores SVM’s proficiency in navigating high-dimensional feature spaces, while both hybrids demonstrated markedly improved generalization over standalone CNN, particularly for noisy or under-represented samples.

These outcomes affirm the efficacy of hybrid paradigm in advancing automated diagnostic systems, effectively mitigating longstanding challenges like feature complexity, data scarcity, and over fitting in medical imaging. By decoupling feature extraction and classification, the framework delivers enhanced flexibility and reliability rendering it particularly apt for resource-limited clinical settings- such as those in southern India where scalable ophthalmic screening remains critically essential.

IX. FUTURE WORK

Future investigations will integrate pre-trained transfer learning architectures (e.g., ResNet-50/101, EfficientNet series and Vision Transformer) to capitalize on extensive prior training and improve feature discriminability through limited fine-tuning We’ll also expand to multi-model datasets from clinicians in TN hospitals, plus global ones like APTOS and EyePACS. That’ll make our models tougher against demographic biases and real-world messiness, such as motion blur or bad lighting.

Interpretability is key for doctors to actually use this in clinics. We’ll integrate explainable AI(XAI) tools like Grad-CAM++,SHAP,and LIME to create heatmaps and attribution maps. These will clearly show ophthalmologists exactly how the model makes its decisions.Edge deployment on mobile devices and API integration with telemedicine platforms will facilitate real time screening in rural outreach programs.

Looking ahead, longitudinal studies could dive into federated learning for privacy- safe training across multiple institutions, paired with ensemble methods that blend hybrid outputs for diagnostic accuracy close to human levels. This work sets a strong foundation for smart ophthalmic systems ready to revolutionize eye care accessibility.

REFERENCES

- [1] Jain, P., Motwani, D., & Sharma, P. (2025). A hybrid adaptive CNN-based framework for eye disease detection using fundus images. *Journal of Information Systems Engineering and Management*, 10(1), 1–15.
- [2] Arulselvam, T., & Joseph, S. J. S. A. (2025). Automated retinal disease detection using anchor-free modified Faster R-CNN. *Journal of Medical Imaging and Health Informatics*, 15(1), 55–67.
- [3] Simanjuntak, R. B. J., Fu'adah, Y., Magdalena, R., Saidah, S., Wiratama, A. B., & Ubaidah, I. D. S. (2025). CNN-based cataract classification using fundus images with multi-class analysis. *International Journal of Computer Vision and Biomedical Imaging*, 12(2), 101–112.
- [4] Tripathi, U., Goel, S. P., Tripathi, J. N., & Madkar, S. R. (2025). Digital image processing-based cataract detection for early diagnosis in elderly populations. *International Journal of Healthcare Informatics and Vision Systems*, 7(1), 14–25.
- [5] Prakash, P. N. S., Sudharson, S., Woonna, V. A., & Bacham, S. V. T. (2025). Cataracts NET: A lightweight deep convolutional neural network for early cataract detection using smartphone images. *EAI Endorsed Transactions on Pervasive Health and Technology*, 11(2), 1–12.
- [6] Mano Aarthi, V. M., Tiwari, N., Khobragade, V., Pareek, M., Panchbhai, A., & Ghosh, P. (2025). Clinical evaluation of an AI-based smartphone cataract screening system in a tertiary care hospital. *Journal of Ophthalmology and Digital Health*, 4(1), 33–44.
- [7] Anjan, N. R., Haral, R., Shejul, A., Harne, K., & Bhat, S. (2025). Deep learning-based mobile application for cataract detection and classification. *International Journal of Mobile Health Applications*, 9(1), 60–71.
- [8] Sharma, H., Wasim, J., & Sharma, P. (2024). An efficient deep learning-based system for ophthalmic disease identification using EfficientNet-B5. *Engineering, Technology & Applied Science Research*, 14(2), 10234–10241.
- [9] Aslam, J., Arshed, M. A., Iqbal, S., & Hasnain, H. M. (2024). Deep learning-based multi-class eye disease classification using pre-trained CNN models. *Technical Journal, University of Engineering and Technology (UET) Taxila*, 29(1), 45–55.
- [10] Dahiya, R., Agarwal, N., Singh, S., Verma, D., & Gupta, S. (2024). Machine learning-based approaches for early detection of diabetic retinopathy. *EAI Endorsed Transactions on Internet of Things*, 10(39), 1–10.
- [11] Rashid, F., Abate, J., & Abdi, A. (2024). Deep convolutional neural network for multiclass classification of external eye diseases. *Journal of Artificial Intelligence in Medical Imaging*, 6(2), 88–97.
- [12] Subin, P. G., & Kannan, P. M. (2024). Hybrid adaptive mutation swarm optimization with regression neural network for eye disease detection. *International Journal of Intelligent Healthcare Systems*, 8(1), 21–33.
- [13] Madhumithra, G., & Sasikala, G. (2022). CNN-based retinal disease classification using fundus images from the STARE database. *International Journal of Research and Analytical Reviews*, 9(3), 456–462.
- [14] Nawaz, A., Ali, T., Mustafa, G., Babar, M., & Qureshi, B. (2022). A memory-efficient convolutional neural network for multi-class retinal disease classification. *IEEE Access*, 10, 118765–118776.
- [15] Lee, Y., Kim, H., & Park, J. (2021). Enhancing iris images using GAN-based reconstruction for improved recognition under challenging conditions. *Journal of Biometrics and Pattern Recognition*, 14(3), 215–227.
- [16] Jayadev, N., & Bellary, S. (2021). Hybrid MDLNN and bacterial foraging optimization for alcohol-related iris damage detection. *International Journal of Intelligent Optical Systems*, 8(2), 95–108.
- [17] Mendes, R. S., Silveira, M. L., Rocha, A. M., & Carvalho, A. C. (2009). Ultrasonographic evaluation of posterior segment abnormalities in patients with mature cataracts. *Clinical Ophthalmology*, 3(1), 1–7.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details