International Research Journal of Multidisciplinary Scope (IRJMS), 2025; 6(3):1045-1066

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2025.v06i03.04649

Enhanced Detection of Retinopathy of Prematurity in Premature Newborns Using Integrated Convolutional Neural Networks and Long Short-Term Memory Architectures

Karkuzhali S*, Selva Madan M, Prathapan C

Department of Computer Science and Engineering, Mepco Schlenk Engineering College, Sivakasi, Tamil Nadu, India. *Corresponding Author's Email: karkuzhali@mepcoeng.ac.in

Abstract

Premature new-borns are at risk for Retinopathy of Prematurity (ROP), a condition that can result in blindness. Those with low birth weights are more likely to experience this. Early detection and treatments are crucial to prevent blindness or severe visual impairment. Unfortunately, the current methods for diagnosing ROP are sometimes subjective, labour-intensive, and require specialized knowledge, which delays the process of both diagnosis and treatment. Deep Learning (DL) algorithms show remarkable efficiency in a variety of medical imaging tasks, including the detection and classification of illnesses. The suggested hybrid Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) model, which offers a viable solution, combines the best aspects of the CNN-LSTM architectures. The CNN component efficiently extracts spatial properties from retinal images, whereas the LSTM component processes sequential image data across time to capture temporal dependencies. The diagnostic precision is further enhanced by the model's increased ability to precisely analyse the progression of ROP. Blood vessel segmentation in fundus image is used to detect the abnormality effectively. Modified MultiResUNet model is used for segmentation of blood vessels. The robustness and accuracy of the suggested model are assessed using performance metrics such the F1-score, sensitivity, specificity, precision, and accuracy. Comparative studies using existing screening methods demonstrate that the proposed deep learning methodology is more effective in terms of accuracy of 97%.

Keywords: Convolutional Neural Network, Deep Learning, Long Short-term Memory, Retinopathy of Prematurity, Segmentation.

Introduction

Retinopathy of Prematurity (ROP) is an eye ailment that primarily affects preterm infants, particularly those with low birth weights and those who need to be in neonatal intensive care units (NICUs) for extensive medical treatment, including oxygen therapy. ROP manifests due to aberrant development of blood vessels in the retina, the light-sensitive tissue essential for vision located at the back of the eye. The effects of ROP can vary widely, encompassing a spectrum from mild to severe. Infants with severe ROP may experience enduring complications such as strabismus (eye misalignment), amblyopia (lazy eye), and diminished visual acuity. Moreover, ROP's impact extends beyond mere vision impairment, influencing the overall quality of life and developmental trajectory of affected individuals. Managing ROP presents challenges due to its intricate nature and the distinct vulnerabilities of premature infants. Conventional diagnostic techniques, like indirect ophthalmoscopy, rely heavily on the expertise of ophthalmologists, introducing subjectivity and variability into diagnosis and treatment decisions. Furthermore, limited access to specialized eye care services, particularly in resource-constrained settings, exacerbates delays in diagnosis and intervention. Every year, ROP affects over 14.38 million babies worldwide, with 23% of those babies being born in India. ROP is more common in women who are less than 32 weeks pregnant or have a birth weight of less than 2000 grams. Approximately 490,000 babies are born in India under 32 weeks of pregnancy each year. ROP incidence and severity are significantly influenced by oxygen administration

(Received 26th March 2025; Accepted 28th June 2025; Published 27th July 2025)

and duration. ROP develops in five phases; zones I through III denote the location, and "plus disease" denotes the advancement of the disease. Better visual outcomes depend on prompt diagnosis and treatment in accordance with the Early Treatment of ROP (ET-ROP) criteria. Automated prediction techniques are desperately needed, given the significant variability and interobserver disagreement in ROP diagnosis, as well as the rising incidence rates.

Numerous research has examined the effectiveness of deep convolutional neural networks (DCNNs) in the classification of ROP. DCNNs have demonstrated potential in automated diagnosis. However, existing methods often face challenges such as computational complexity or limited accuracy. By utilizing the advantages of both architectures, the proposed hybrid model that combines CNN and LSTM offers a comprehensive method for detecting ROP. CNNs are well-suited for extracting spatial features from retinal images, which is crucial for identifying pathological changes indicative of ROP. In the context of ROP detection, the CNN component of the hybrid model analyzes retinal images to capture intricate patterns, textures, and structural abnormalities associated with the disease. Through a series of convolutional layers followed by pooling operations, the CNN efficiently learns hierarchical representations of the input images, effectively capturing relevant features for ROP diagnosis. However, retinal image data is inherently sequential, capturing the progression of ROP over time. This temporal aspect necessitates the integration of recurrent neural networks (RNNs), such as LSTMs, to capture temporal dependencies within the sequential image data. LSTMs excel at modeling sequential data by preserving long-range dependencies and selectively retaining relevant information over time through a series of memory cells and gates. In the CNN-LSTM hybrid model for ROP detection, the LSTM component processes the sequential retinal image data generated over time. By sequentially feeding retinal images into the LSTM network, the model learns to capture temporal patterns and changes indicative of ROP progression. This enables the model to discern subtle alterations in retinal morphology and pathology, facilitating more accurate detection of ROP.

With its extensive collection of fundus images of premature infants, the HVDROPDB dataset constitutes a major contribution to the field of ROP research. The Ret Cam and Neo imaging systems were used to obtain posterior and temporal views of these images, which provide important information on the onset and course of ROP. Taking its name from the prestigious H.V. Desai Eye Hospital in Pune, India, where the data was collected, the dataset consists of photographs that were acquired by skilled optometrists through laborious screening procedures. High-resolution photos with a 120° field of view were taken using the Ret Cam and Neo imaging systems, which are popular and easily accessible. This allowed for thorough inspection of the retinal structures. There are other subsets within the dataset that are specific to optics, such as HVDROPDB-OD, HVDROPDB-BV and HVDROPDB-RIDGE. These subsets make it easier to segment and analyze retinal structures in-depth, which is necessary for identifying ROP zones and phases. These subsets make it easier to segment and analyze retinal structures in-depth, which is essential for identifying ROP phases and zones.

Retinal images of babies with birth weights of 3000 g or less and gestational ages between 26 and 36 weeks are included in the dataset. Pupils were dilated using a standardized eye drop approach prior to picture acquisition, guaranteeing consistent and high-quality images for all patients. Every baby's eye was painstakingly studied from the nasal, superior, inferior, and temporal perspectives, yielding an extensive collection of about 8,280 Neo and 10,570 RetCam photos from 1,100 patients. Researchers and physicians investigating ROP can find a plethora of information from the photos, which are saved in PNG and JPEG formats, respectively. It is an important tool for furthering ROP research and makes the creation and verification of automated ROP screening systems easier. It does this by giving users access to a variety of carefully captioned photos. The incidence of respiratory failure within the first 72 hours of life was defined as the need for endotracheal surfactant or mechanical ventilation. Nebulized poractant alfa was administered along with nasal continuous positive airway pressure (NCPAP) in preterm infants with respiratory distress syndrome (RDS) to compare its effectiveness and safety with

NCPAP (1). A shift from traditional machine learning approaches to deep convolutional neural networks (DCNNs) for artificial intelligence (AI) applications in ROP was emphasized. Generalizability of algorithms and integration into clinical workflows were identified as the main barriers to AI use in ROP detection and treatment (2).

An approach for identifying Aggressive Posterior ROP (AP-ROP), a rapidly progressing form of ROP leading to childhood blindness, was developed. An end-to-end automatic diagnosis solution was proposed using two distinct networks, incorporating channel attention for enhanced feature representation (3). The performance of a deep learning algorithm for ROP screening in Mongolia and Nepal was assessed. Retrospective analysis was conducted on clinical data collected prospectively, using fundus images acquired with different camera systems (4). The prevalence and risk factors for ROP and severe ROP in preterm newborns in Northern India were investigated. Among 340 screened infants, 63 (18.5%) were diagnosed with ROP, with 8 (2.4%) presenting with severe cases. Stages of ROP were reported as 3.5% in stage 3, 30.2% in stage 1, and 63.5% in stage 2 (5).

The therapeutic usefulness of systemic inflammatory indices in predicting moderate to severe bronchopulmonary dysplasia (BPD) in preterm infants was examined. At birth and at BPD diagnosis, indices such as SII, PIV, MLR, NLR, PLR, and SIRI were measured (6). Clinical features and causative factors for asymmetrical ROP severity in premature twin pairs were analyzed. It was found that prolonged oxygen supply and higher frequency of blood transfusion correlated with increased ROP severity (7). At two-year follow-up, the median spherical equivalent (SE) was -0.13 D (IQR, 4.20 D). Amblyopia, nystagmus, strabismus, optic atrophy, and high myopia were the most common ocular findings. Favorable structural outcomes and reduced progression to retinal detachment were observed in neonates treated with intravitreal ranibizumab and delayed laser therapy (8). Machine learning techniques were applied to classify neonatal mortality and morbidity. It was concluded that low birth weight and gestational age increase mortality risk, and long-term morbidities such as BPD, NEC, and ROP were predicted using ML models based on low-birth-weight infants (9).

The role of gut microbiota in the immune system development of premature infants was highlighted. Dysbiosis was identified as a common condition impairing immune maturation and contributing to mortality in children under five. ROP was recognized as a condition impeding retinal blood vessel development, leading to blindness (10). A discrete conditional phase-type model combined with support vector machine (SVM) was proposed for ROP prediction. Clinical data were modeled using conditional phase-type distributions and integrated with an SVM classifier to predict severity levels, addressing class imbalance using a novel decomposition technique (11). The link between systemic inflammation and encephalopathy of prematurity (EoP), associated with neurocognitive impairment, was established (12). A training methodology using pairwise comparisons of image patches, rather than direct class labels, was proposed to improve deep learning model efficiency for ROP detection (13). The use of digital systems for ROP diagnosis was emphasized. Advanced image analysis and deep learning were employed to enhance screening accuracy, efficiency, and accessibility (14). An automated approach for ROP screening was implemented using B-COSFIRE filters for initial segmentation, followed by optic disc removal and binarization. Tortuosity estimation was used for ROP assessment, achieving 88% sensitivity and 94% specificity (15).

Anesthetic methods and intraoperative care strategies for premature infants weighing under 1.5 kg undergoing TC-PDA were investigated. Only preterm infants with significant PDA undergoing closure at a specific institution were included (16). The genotype-phenotype correlations of a microdeletion on the X chromosome causing Norrie disease and X-linked Kabuki syndrome were studied. Insights into the inheritance patterns and overlapping clinical features were provided, underscoring the importance of molecular diagnosis (17). Retinal blood vessel extraction from images was performed using a matched filter and first-order difference of Gaussians (FDOG), with thresholding based on the filter responses (18).

Surveillance of congenital rubella syndrome (CRS) was conducted using data from 14 sentinel sites across India between 2016 and 2021, with 3940 probable CRS cases enrolled (19). The importance of retinal blood vessels in diagnosing retinal disorders was emphasized through a model automating vessel segmentation and abnormality classification, evaluated on three benchmark datasets using performance metrics (20).

In this work, developed an automated technique for ROP prediction in order to overcome the substantial variability and discrepancies between observers in ROP diagnosis. Using the combined Modified MultiResUNet with a matching model to isolate the retinal vasculature from the fundus images. The contour features and gray level co-occurrence matrix (GLCM) of segmented images are extracted and selected through the use of an embedded feature selection technique. Following its evaluation using permutation significance, the selected features are classified using the Random Forest classifier. ROP is a proliferative vitreoretinopathy condition that affects premature babies. If left untreated, it frequently leads to irreversible blindness. The survival rates of preterm infants have significantly increased due to breakthroughs in neonatal medical treatment, which has raised the prevalence of ROP globally.

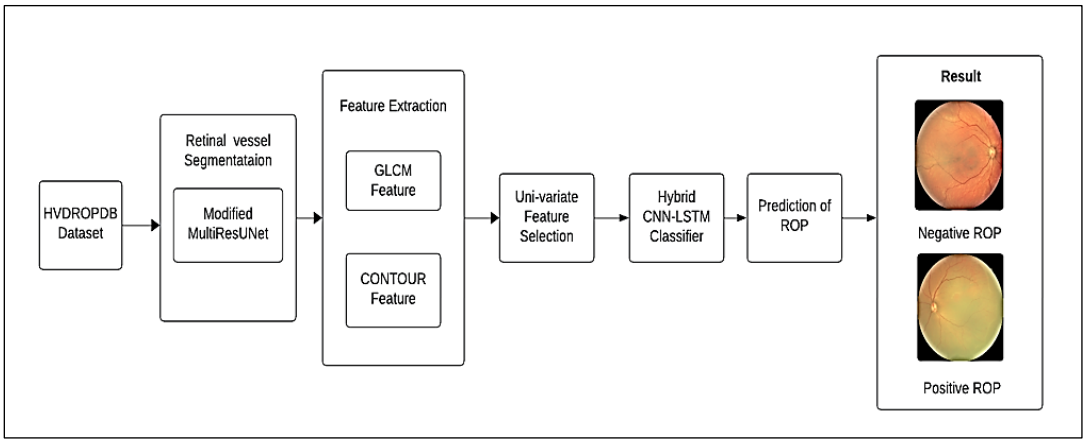

Figure 1: System Architecture for ROP Detection

Methodology

For effective detection of ROP in infants, the research work proposes a deep learning-based model with hybrid CNN-LSTM networks. The research work involves the fundus image dataset images are used, segmentation of collected data, feature extraction and selection and finally classifying then into ROP or not ROP. The system architecture of the proposed work is depicted in Figure 1.

HVDROPDB Data-Set

A vital tool for academics working to advance ROP screening system automation is the HVDROPDB Ret Cam Neo Segmentation dataset. This dataset provides a thorough basis for algorithmic development. It consists of retinal fundus images of premature newborns acquired using Ret Cam and Neo imaging systems at PBMA's H.V. Desai Eye Hospital, Pune. In order to create interpretable algorithms that will automate ROP screening procedures, skilled annotators carefully marked out ground realities for the optic disc, blood vessels, and demarcation line/ridge structures in the photographs. There are 600 images are present in the HVDROPDB Dataset. This study clarifies the dataset's structural integrity, annotation process, and intrinsic importance in propelling ROP screening technology developments. Interestingly, the study included babies whose gestational ages less than 32 weeks pregnant or have a birth weight of less than 2000 grams. Approximately 490,000 babies are born in India under 32 weeks of pregnancy each year. ROP incidence and severity are significantly influenced by oxygen administration and duration. Figure 2 is the fundus image captured using Neo Imaging and the ground truth mask for that image.

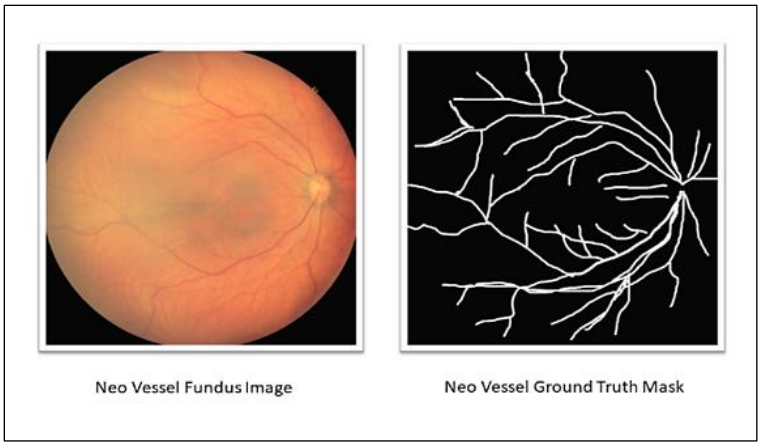

Figure 2: Neo Vessel Fundus Image and Its Mask

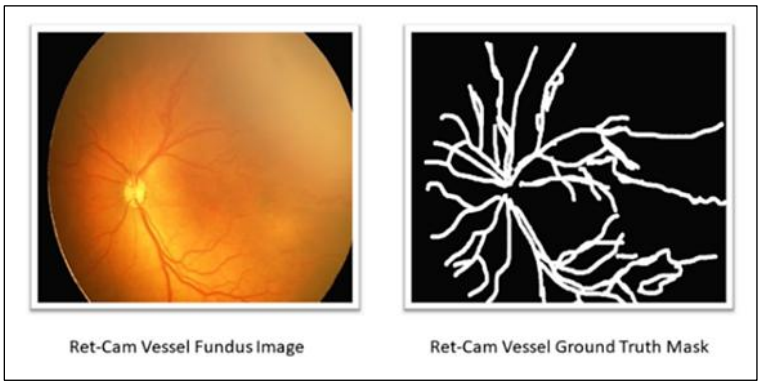

Figure 3: Ret-Cam Fundus Image and Its Mask

Table 1: HVDROPDB Dataset Description Image Size 506 x 506

Dataset Name	Sub-Dataset		Number of Images
HVDROPDB-OD	RetCam	images	50
		masks	50
	NeoCam	images	50
		masks	50
HVDROPDB-BV	RetCam	images	50
		masks	50
	NeoCam	images	50
		masks	50
HVDROPDB-RIDGE	RetCam	images	50
		masks	50
	NeoCam	images	50
		masks	50

Figure 3 depicts the fundus image taken using Ret-cam is a wide-angle fundus camera and the ground truth mask for the same. Utilizing Ret Cam and Neo devices, images were captured from various perspectives, including posterior, temporal, superior, inferior, and nasal views of both left and right eyes, resulting in an average of 2 to 12 images per eye. Table 1 gives the detailed description of the dataset used and the number of images present in each sub dataset.

Retinal Vessel Segmentation

The proposed Modified MultiResUNet architecture presented in this study serves as an efficient framework for semantic segmentation tasks, particularly tailored for medical image analysis, with a focus on retinal structure segmentation in the context of Retinopathy of Prematurity (ROP) screening. The network architecture comprises encoding and decoding stages, leveraging convolutional and transposed convolutional layers for feature extraction and spatial up sampling, respectively. Key modifications and enhancements in the proposed Modified MultiResUNet include:

Multi-Resolution Blocks: These blocks enable the network to capture and process features at multiple scales, which is critical for accurately segmenting retinal structures that vary in size and appearance.

ResPath Connections: These connections provide residual links that help in preserving spatial information across different levels of the network, reducing the vanishing gradient problem, and improving the gradient flow during training.

Improved Skip Connections: By refining the skip connections, the model ensures better feature propagation and merging of high-resolution spatial information with low-resolution contextual information, leading to more precise segmentation boundaries.

Efficient Use of Parameters: The architecture is designed to be parameter-efficient, ensuring that it can be trained effectively even with limited medical image datasets, which is often a constraint in clinical applications. These enhancements make the Modified MultiResUNet particularly suitable for the challenging task of ROP screening, where accurate segmentation of retinal structures is essential for early diagnosis and treatment. The ability to capture fine-grained details and contextual information simultaneously ensures that the model can reliably segment the intricate retinal features associated with ROP. This contributes significantly to improving the accuracy and reliability of ROP screening, potentially leading to better clinical outcomes for affected infants. At the encoding stage, a series of convolutional blocks followed by max-pooling operations progressively down sample the input image, enabling the extraction of hierarchical feature representations. The encoding pathway consists of four encoder blocks, each comprising convolutional layers and batch normalization, facilitating the extraction of intricate spatial features while preserving spatial resolution. Figure 4 shows the Retinal Vessel Segmentation using Modified MultiResUNet.

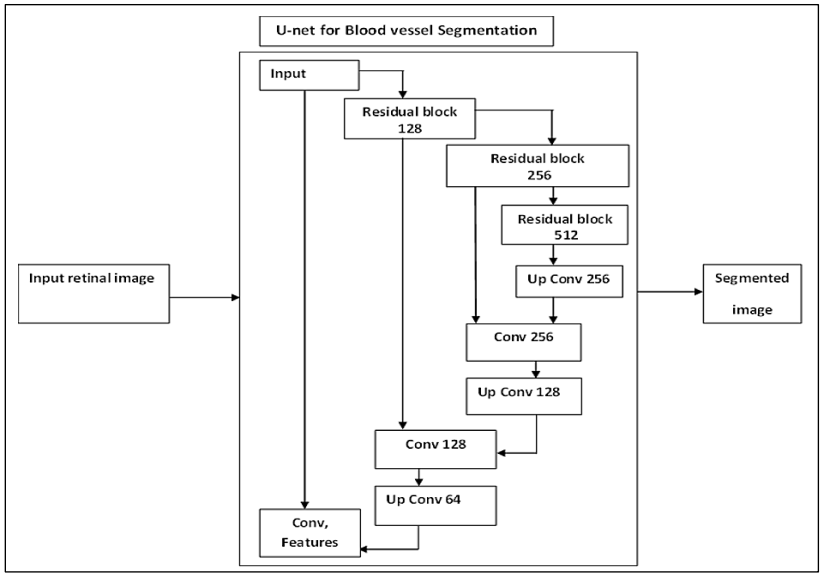

Figure 4: Retinal Vessel Segmentation using Modified MultiResUNet

Several important criteria are used to assess the effectiveness of the Retcam segmentation and neo-vessel segmentation algorithms. Accurate segmentation is shown by the Jaccard Index, which

measures the overlap between anticipated and ground truth masks. Preciseness and recall are balanced in the F1 Score.

Each decoder block receives inputs from the bottleneck layer and the corresponding encoder block through skip connections, enabling the incorporation of both high-level semantic information and low-level spatial details. The decoder blocks refine feature representations at multiple scales, progressively up sampling the spatial resolution while enriching feature maps with contextual information. Finally, the classifier module employs a single convolutional layer.

The bottleneck layer acts as a feature fusion mechanism, integrating high-level semantic information from the encoding stage. This layer enhances the network's discriminative power by consolidating abstract features learned from the encoding pathway. The decoding pathway employs transposed convolutional layers to facilitate up sampling and reinstatement of spatial resolution.

The output segmentation map represents the network's prediction for the presence of relevant anatomical structures. The proposed MultiResUNet architecture exhibits a balanced trade-off between receptive field size, feature resolution, and computational efficiency, making it well-suited for automated ROP screening systems, with potential applications in clinical settings to assist healthcare professionals in early detection and management of ROP.

The Dice loss, an integral component in segmentation tasks, quantifies the similarity between predicted and ground truth masks by measuring their overlap. It calculates the intersection over union (IoU) of the two masks, providing a measure of segmentation accuracy ranging from 0 to 1. On the other hand, the Dice binary cross-entropy (BCE) loss combines the Dice loss with BCE to create a hybrid loss function. This fusion optimizes both segmentation accuracy and binary classification performance, offering a balanced approach to training segmentation models. Widely utilized in medical imaging and computer vision it enhances model robustness and performance across various applications. The dataset is divided into training and validation sets, and data loaders are utilized to efficiently load and preprocess the data for training.

Convolutional and pooling layers are used in the construction of the Multi U-Net model for both encoding and decoding, along with skip connections to maintain spatial information. During the training phase, an Adam optimizer with a predetermined learning rate is used for optimization, and a Reduce-LR-On-Plateau scheduler is used to dynamically modify the learning rate. To maximize segmentation accuracy, a mix of binary cross-entropy loss and dice loss is used as the loss function. To guarantee that the best-performing model is kept throughout training, validation loss improvement is used to save model checkpoints. Overall, the code shows how to train a Multi U-Net model for retinal image segmentation using a thorough pipeline that includes features for effective data management and model tuning.

The application of an automated semantic segmentation model to retinal fundus images in order to identify ROP. The model is trained and assessed on a collection of images taken with several imaging devices, including RetCam and Neo, and annotated by ROP specialists, using the MultiResUNet architecture. Inference is carried out on a test dataset, a pre-trained MultiResUNet model checkpoint is loaded, and measures like the Jaccard Index, F1-score, Recall, Precision, and Accuracy are used to assess the model's performance. The outcomes are stored for later examination, together with the original photos, ground truth masks, and forecasted masks. The script also computes frames per second (FPS), for the purpose of assessing inference speed. This automated method seeks to support the creation of an understandable and effective ROP screening system, enhancing visual.

Feature Extraction

The GLCM and contour features of segmented images are extracted and selected using an embedded feature selection method. The extraction of six GLCM characteristics of the pixel with its neighbour, including dissimilarity, homogeneity, contrast, Angular Second Moment (ASM), correlation, and energy at 0˚, 45˚, 90˚, and 135˚. The contour features used are area, perimeter, two centroids, and 24 moments of the retinal vessels segmented. These

features are fused to develop a 52-length feature vector.

Dissimilarity: Dissimilarity represents the difference in intensity between pixels at various positions and orientations within the image, quantified through the GLCM method. Computed at angles of 0˚, 45˚, 90˚, and 135˚, Dissimilarity provides insights into the level of heterogeneity or dissimilarity in pixel intensities across different orientations. A higher value indicates greater dissimilarity, implying a more complex texture with varying pixel intensities, while a lower value suggests more uniform texture.

$$Dissimilarity = \sum_{x,y=0}^{n-1} I_{xy}|x-y| \quad ...[1]$$

Homogeneity: Homogeneity, computed using GLCM, measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. It characterizes the uniformity or smoothness of the texture within the image. Higher homogeneity values signify that the pixel intensities are closely packed around the GLCM diagonal, indicating a more homogeneous texture. Conversely, lower homogeneity values indicate a more heterogeneous texture with varying pixel intensities distributed away from the GLCM diagonal.

$$Homogeneity = \sum_{x,y=0}^{n-1} \frac{I_{xy}}{1+(x-y)^2} \quad ...[2]$$

Contrast: Contrast, derived from the GLCM, quantifies the local variations in intensity levels within the image. It measures the difference in intensity between neighbouring pixels, emphasizing the sharpness of transitions between different texture regions. Higher contrast values indicate stronger variations in pixel intensities, suggesting a more textured or detailed image, whereas lower values signify smoother transitions between pixel intensities and a less textured appearance.

$$Contrast = \sum_{x,y=0}^{n-1} I_{x,y}(x-y)^2 \quad ...[3]$$

Angular Second Moment: ASM is also known as Energy, reflects the overall uniformity or homogeneity of the image texture. Computed from the GLCM, ASM represents the sum of squared elements in the GLCM, emphasizing the dominance of certain pixel pairs in the texture distribution. Higher ASM values indicate a more uniform texture with a balanced distribution of pixel pairs, whereas lower values suggest a more heterogeneous texture with a skewed distribution of pixel pairs.

$$ASM = \sum_{x,y=0}^{n-1} I_{xy}^2 \quad ...[4]$$

Correlation: Correlation, computed from the GLCM, describes the linear dependency between pixel intensities at different image locations and orientations. It measures the degree to which pixel pairs in the image exhibit linear correlation in their intensities. Correlation values range from -1 to 1, where values closer to 1 indicate stronger positive correlation, implying a more ordered or structured texture, while values closer to -1 suggest stronger negative correlation, indicating a more disordered or random texture.

$$Correlation = \sum_{x,y=0}^{n-1} \frac{(x-m_x)(y-m_y)}{\sqrt{(n_x^2)(n_y^2)}} \quad ...[5]$$

Energy: Energy, derived from the GLCM, represents the sum of squared elements in the GLCM. It quantifies the overall uniformity or homogeneity of the image texture, emphasizing the dominance of certain pixel pairs in the texture distribution. Higher energy values indicate a more uniform texture with a balanced distribution of pixel pairs, whereas lower values suggest a more heterogeneous texture with a skewed distribution of pixel pairs.

$$Energy = \sqrt{ASM} \quad ...[6]$$

Area: The Area feature represents the total area of the image, computed as the sum of pixel intensities. It provides a measure of the spatial extent covered by the image, indicating its size or magnitude.

Contour Perimeter: Contour Perimeter refers to the cumulative perimeter of all contours detected within the image. It quantifies the total length of the boundaries of distinct regions or objects present in the image, providing insights into the complexity or irregularity of the image's shape or structure.

Centroid: The Centroid coordinates denote the geometric center of the contours detected within the image. They represent the average position of the contour pixels along the X and Y axes, providing a reference point for the spatial distribution of objects or regions within the image.

Traversing a directory containing segmented images to extract features, while simultaneously collecting image paths and inferring labels from filenames, facilitating subsequent analysis or machine learning endeavours. Following feature extraction, the shape of the resultant feature matrix (X) and label array (y) is printed, providing valuable insights into the dimensions of the dataset and its preparedness for further investigation. This structured approach ensures systematic feature extraction and dataset preparation, laying a solid foundation for subsequent analysis or model training activities.

Feature Selection: The best features for the machine learning classifier are chosen using the embedded feature selection approach based on ANOVA-F statistics. Using the univariate feature selection method with k = 10 and the ANOVA-F value as the scoring scale, the 52-length feature vector is reduced to the ten best features. The ten features are as follows: energy, perimeter, area, centroid, central moments, contrast, ASM, correlation, scale invariants, rotation invariants, and dissimilarity. The univariate approach uses a predetermined scoring system to evaluate each characteristic separately in order to identify the most important qualities classifier's performance, based on how important each individual feature is to the it chooses the top 10 features.

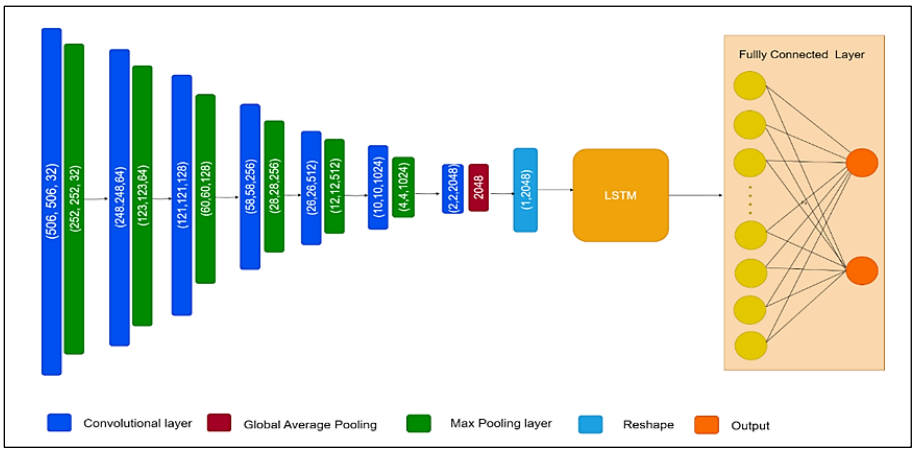

Figure 5: Architecture of CNN-LSTM Model

The highest F-Statistic value of the ANOVA between the particular feature and the output label to determine the presence of ROP. Initially, the dataset is partitioned into training and testing subsets, facilitating subsequent evaluation. Feature selection is executed utilizing the ANOVA-F statistic method, with the Select K-Best function extracting the most salient features. These selected features are then employed to train a Random Forest classifier, configured with a specified number of estimators and random state. Subsequently, the classifier's performance is evaluated on the test set, yielding an accuracy metric indicative of its predictive efficiency. Furthermore, permutation importance is computed to discern the relative. The importance values obtained are arranged in a way that emphasizes the most significant characteristics, providing insightful information about the relevance of the features and supporting the interpretation and improvement of the model. The methodical path from dataset preparation to model evaluation, which is necessary

for reliable machine learning model building, is highlighted by this structured methodology.

Classification: To detect and classify the ROP affected and ROP not affected images, used a hybrid CNN-LSTM model. As a part of the hybrid model, the main objective of CNN is to extract the relevant features from the input images. To do so, we incorporate multiple convolutional layers for the creation of CNN with pooling layers gives feature map as the result. Next the LSTM processes the extracted features in a sequential manner, learns the patterns in the features. Finally, the output of the LSTM is fed into a fully connected layer for classification. The developed model is then trained on a labelled dataset. The trained model is then evaluated. For more effective classification the fine-tuning is done by changing the hyper parameters of the model such as learning rate, batch size, optimizer. Figure 5 shows the architecture of CNN-LSTM model. The proposed hybrid model architecture is summarized in Table 2. It begins with an input layer accepting fundus images of size 506 × 506 × 3. The spatial features are extracted using a pre-trained ResNet50 backbone (excluding its top classification layer), known for its deep residual connections and efficiency in medical imaging tasks. A global average pooling layer follows to reduce the spatial dimensions and produce a compact feature vector. This vector is reshaped into a sequence format compatible with the LSTM layer, which consists of 128 units to capture temporal dependencies within the extracted features. To prevent overfitting, a dropout layer with a rate of 0.3 is included. The LSTM output is then passed through a dense layer with 64 units and ReLU activation. Finally, a softmax-activated output layer classifies the image as ROP-affected or not affected.

To enhance the model's generalization performance and address the challenge of limited data, various data augmentation techniques were applied to the training images. These included horizontal and vertical flipping, random rotations, mirroring, and zooming. The goal was to introduce meaningful variability in image presentation while preserving essential anatomical structures relevant to ROP detection. This augmentation helped the model learn robust spatial features that are invariant to position and orientation, reducing the risk of overfitting. Additionally, although not a direct substitute for formal class rebalancing, these augmentation techniques also contributed to improving class representation for underrepresented ROP stages, thereby supporting more balanced learning across categories.

- **Horizontal Flipping**: Simulates natural variation in eye orientation to improve generalization.
- **Vertical Flipping**: Introduces vertical orientation diversity to reduce bias toward fixed patterns.
- **Rotation (±10–30°)**: Mimics image capture variation, enhancing rotational invariance.
- **Mirroring**: Adds geometric variability, increasing dataset diversity.
- **Random Zoom (90–110%)**: Simulates different image scales, improving robustness to size variations.

In the proposed ResNet50–LSTM hybrid model, the architecture begins with a pre-trained ResNet50 backbone (50 layers) to extract deep spatial features from the input fundus images. These features are passed through a Global Average Pooling layer and reshaped for sequential processing. A Long Short-Term Memory (LSTM) layer with 128 units is employed to capture temporal dependencies in the feature sequence. This is followed by a Dense layer with 64 neurons using the ReLU activation function, and a final output layer with 2 neurons using Softmax activation for binary classification. To prevent overfitting, a dropout layer with a rate of 0.3 is included after the LSTM layer. This architecture effectively integrates spatial and temporal learning for robust ROP classification.

Although the input to the model consists of static fundus images, the extracted spatial features often exhibit sequential patterns that reflect the temporal progression of ROP. The LSTM layer is incorporated to capture these temporal dependencies in the feature space, enabling the model to better recognize progression cues such as vessel dilation, ridge formation, and neovascularization, which evolve over time

To make efficient detection of affected eye, the proposed work suggests the use of hybrid CNN-LSTM neural network model. This consists of a CNN

of seven layers each followed by a max pooling layer and an LSTM layer which ultimately gives the detected result. This neural network architecture uses a convolutional and recurrent layer combination that is specifically designed for image classification.

First, a series of convolutional layers are used, and then max-pooling layers are added with the intention of extracting hierarchical features from input images. These convolutional layers can capture a wide range of features, from simple to complicated features. A max-pooling layer, which reduces spatial dimensions while keeping important characteristics, which comes after each convolutional layer. To add non-linearity, ReLU activation functions are used inside the convolutional layers. A global average pooling layer is used after the convolutional layers to reduce the number of spatial dimensions to a single vector. The LSTM layer, a kind of RNNskilled at processing sequential data, is then accommodated by reshaping this vector. With the help of the LSTM layer, temporal dependencies in the feature representations produced by the convolutional layers are captured. Incorporating softmax activation into a thick layer allows for the last step of classifying inputs into two groups. Convolutional and recurrent layers are smoothly integrated into this architecture, which effectively learns both spatial and temporal features that are critical for image classification tasks.

The developed model takes image of size (506,506,32) as input which is then given to the subsequent convolutional layers. The convolutional layers have progressively increased filter sizes and feature maps from 32 to 2048, capturing both low-level and high-level features. The model creation algorithm is described in algorithm 1 to algorithm 5. Several tests were conducted to ascertain the ideal regularization hyper-parameter while taking the posterior distribution of the dropout approach into account. The mathematical probability (P) for the dropout procedure ranges from 0.1 to 0.5. The dropout value was started with a lower dropout probability and kept raising it to restrict the transmission of that loss to the subsequent layers.

Table 2: Summary of the ResNet50–LSTM Hybrid Model Architecture

Layer Type	Details
Input Layer	Input fundus image (506 × 506 × 3)
ResNet50 Backbone	Pre-trained on ImageNet (excluding top layer), used for feature extraction
Global Average Pooling	Reduces spatial dimensions to a 1D feature vector
Reshape Layer	Reshapes feature vector to a sequence format suitable for LSTM input
LSTM Layer	128 units, captures temporal dependencies in feature sequence
Dropout Layer	Dropout rate = 0.3, used for regularization
Dense Layer	Fully connected layer with 64 units and ReLU activation
Output Layer	Dense layer with 2 units and Softmax activation for binary classification

Algorithm: Hybrid CNN-LSTM model creation

```
Begin
  Step 1: Initialize the Sequential Model
  model = Sequential()
  Step 2: Add Convolutional and MaxPooling Layers
  for i = 1 to 7
    Add a convolutional layer with specified parameters
model.add(Conv2D(filters, kernel_size, strides, padding, activation))
    Add a max pooling layer with specified parameters
model.add(MaxPooling2D(pool_size, strides, padding))
  end for
  Step 3: Reshape the Output for LSTM Layer
  Reshape the output of the previous layers to the required shape for LSTM
```

model.add(Reshape(target_shape))
Step 4: Add LSTM Layer
Add an LSTM layer with specified parameters
model.add(LSTM(units, activation, return_sequences, dropout))
Step 5: Add Dense Layer
Add a dense layer with specified parameters for the final output
model.add(Dense(units, activation))
Stop

Algorithm: ROP Detection

Input: Raw retinal image data
1. Preprocess input data
   - ip = PreprocessImageData(input data)
2. Segment retinal structures
   - Sip = SegmentRetina(modifiedMultiResUNet, ip)
3. Extract texture features using GLCM
   - f1 = ExtractGLCMFeatures(Sip)
4. Extract contour features
   - f2 = ExtractContourFeatures(Sip)
5. Combine features
   - f = CombineFeatures(f1, f2)
6. Classify using CNN-LSTM model
   - result = CNN_LSTM_Classification(f)
7. Determine result
   - if result < threshold then
     Output: "Normal Retina"
    else
     Output: "ROP Retina"
    end if

Results and Discussion

Evaluation Metrics

The performance evaluation leverages various performance metrics to assess the effectiveness of the proposed approach in predicting ROP. The evaluation of the proposed approach against machine learning classifiers and pre-trained networks, as well as the assessment of ROP-specific features, contributes to the development of a comprehensive framework for automated ROP diagnosis. Dice loss is a function that adapts to the situation where the data being compared looks similar. Dice loss is a function that adapts to the situation where the data being compared looks similar. If not similar, perform a loss. Dice loss is calculated by using the dice coefficient using y true and y pred. Below Equation [7] shows the representation of dice coefficient with X and Y as y true and y pred respectively.

$$Dice\ loss = \frac{2*|X \cup Y|}{|X|+|Y|} \quad \ldots[7]$$

With the above calculated dice coefficient value, the probability of dice loss is given by below Equation [8].

$$\text{Dice loss} = 1 - \text{Dice coef} \quad \ldots [8]$$

Accuracy in segmentation results checks for the predicted output and ground truth's matrix values that enhances the segmentation for the manual separation and automated segmented results which tells higher the accuracy of the results higher the performance and better the segmentation output.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{... [9]}$$

Precision is useful in assessing the proportion of positive identifications in which what all positive proportions are correct. It can be used to evaluate how well a model is able to accurately identify pixels that belong to the target class (positive class) without identifying pixels that do not belong to the class (false positives).

$$Precision = \frac{TP}{TP+FP} \quad \text{...[10]}$$

Sensitivity identifies the proportion of 'actual positive' values that were identified correctly. Which means it helps in identifying how well a model can identify all the relevant pixels i.e., the positive values that are concerned. Sensitivity considers all the relevant pixels in an image, including those that are difficult to identify.

$$Sensitivity = \frac{TN}{TN+FP} \quad \text{...[11]}$$

Specificity refers to the negative values that are correctly predicted. Specificity is a performance metric used in machine learning to evaluate the ability of a model to correctly identify negative samples, i.e., samples that do not belong to a particular class. It measures the proportion of actual negatives that are correctly identified by the model.

$$Specificity = \frac{TP}{TP+FP} \quad \text{...[12]}$$

In the above-mentioned formulae, TP denotes the true positives, TN is for true negatives, FP denotes false positives and FN denotes false negatives.

Retinal Vessel Segmentation

The fundus images in the HYDROPDB dataset are resized before getting into the segmentation. The images are resized to 512x 512 dimension so that it will be compatible with the dimension of the developed MultiResUNet model. Resizing the images to match this expected size ensures that the model can process the images without encountering input dimension mismatches or errors. Resizing fundus images to 512x512 dimensions before using them with the MultiResUNet model is done primarily for standardization, compatibility, and efficiency purposes. This preprocessing step contributes to the overall effectiveness and reliability of the segmentation process in medical image analysis tasks. Using a consistent input size can optimize the computational efficiency of the segmentation model. It allows for better utilization of hardware resources like GPU memory and facilitates the use of batch processing techniques during training, which can speed up the learning process. While resizing can alter the original aspect ratio and possibly distort the image slightly, the aim is to minimize such effects while ensuring that important features and structures in the fundus images are preserved. Techniques like interpolation e.g., bilinear or bicubic) can be employed during resizing to maintain image quality as much as possible. Resizing images to a standard size can positively impact the segmentation performance of the model. When all input images are of the same dimensions, the model learns to extract features and patterns consistently across the dataset, potentially improving the overall accuracy and robustness of the segmentation predictions. Standardizing image dimensions simplifies the training process by reducing the need for complex input handling within the model architecture. It also facilitates easier experimentation with hyperparameters and model configurations, contributing to more efficient model development and tuning.

In Figure 6(A) represents the original fundus image of normal eye, Figure 6(B) is the ground truth which is used for evaluating the segmented image and Figure 6(C) is the result of segmentation using MultiResUNet.

Figure 6: Segmentation of Normal Eye, A) Original Fundus Image, B) Mask of Blood Vessel, C) Segmented Blood Vessels

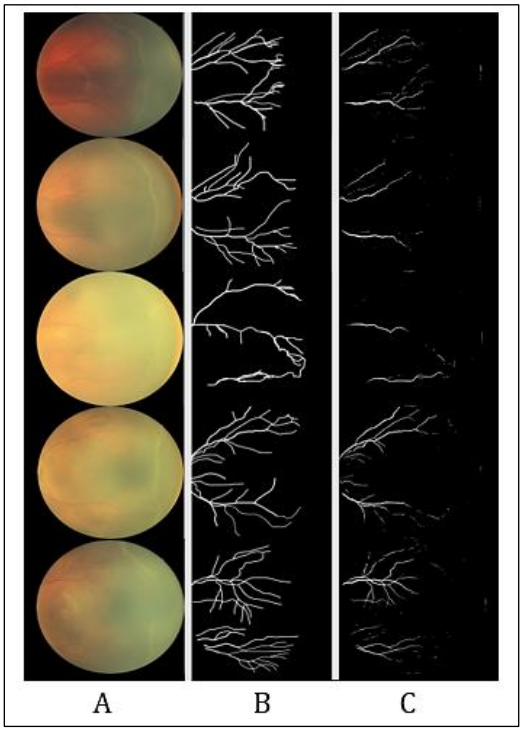

Figure 7: Segmentation of an ROP Eye, A) Original Fundus Image, B) Mask of Blood Vessels, C) Segmented Blood Vessel Image

Figure 7 shows the result of segmentation in an ROP affected eye. Here (A) represents the original fundus image of normal eye, (B) is the ground truth which is used for evaluating the segmented image and (C) is the result of segmentation using MultiResUNet. The created MultiResUNet model is then evaluated using Jaccard, Precision, Recall, Accuracy, F1-Score and FPS. Table 3 and Table 4 tabulates the evaluation results of the MultiResUNet model for Ret-cam dataset and Neo Vessel dataset.

Table 3: Performance Metrics for Ret-cam Segmentation

Performance Metrics	Value
Jaccard coefficient	0.1862
F1	0.3093
Recall	0.2118
Precision	0.6331
Accuracy	0.9710

Table 4: Performance Metrics for Neo Vessel Segmentation

Performance Metrics	Value
Jaccard coefficient	0.1391
F1	0.2348
Recall	0.1498
Precision	0.6827
Accuracy	0.9764

The performance of the Modified MultiResUNet model is assessed by running it on various datasets. The model is assessed in relation to Jaccard, F1-Score, Accuracy, Precision, and Recall. Figure 8 shows the model evaluation mentioned before.

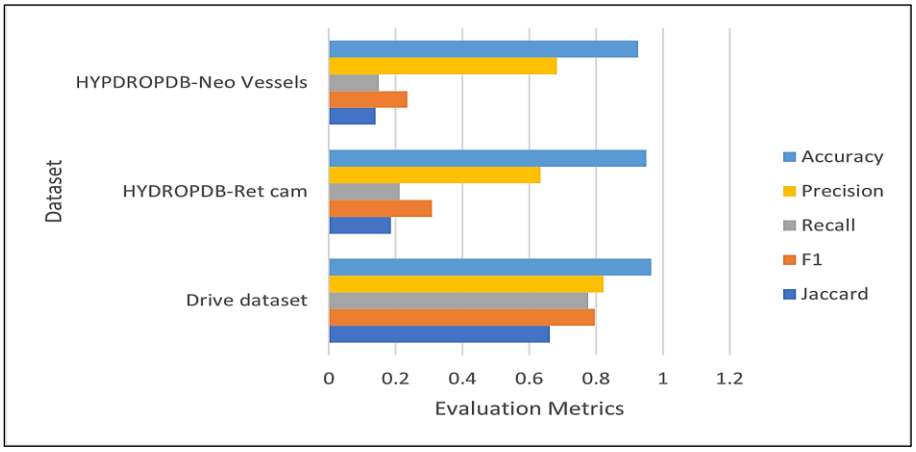

Figure 8: MultiResUNet Algorithm Performance Across Different Datasets

Feature Extraction and Selection

The GLCM features like energy, dissimilarity, ASM, Correlation, Contrast, Homogeneity at 0˚, 45˚, 90˚, and 135˚ and the contour features like centroid, area and parameter are tabulated in Table 5 and Table 6. Feature values describe characteristics of a Dissimilarity, Homogeneity, Contrast, ASM, Correlation and Energy or region within an image, likely related to a segmentation task such as RETCAM or neo-vessel analysis.

Table 5: GCLM and Contour Features

Feature name and Orientation		Value
Dissimilarity	0˚	8.90051675636008

	45˚	12.933544218963622
	90˚	9.645150440313111
	135˚	12.997997097131215
Homogeneity	0˚	0.9650965494898684
	45˚	0.9492809987414923
	90˚	0.9621764623030811
	135˚	0.9490282462435263
Contrast	0˚	2269.6317728718204
	45˚	3298.0537758357236
	90˚	2459.5133622798435
	135˚	3314.4892597684598
ASM	0˚	0.7845505931618361
	45˚	0.7697733911688727
	90˚	0.7818428367462725
	135˚	0.7695463381668028
Correlation	0˚	0.8079705366614376
	45˚	0.7214412916580394
	90˚	0.7919049968026267
	135˚	0.7200531253373984
Energy	0˚	0.8857486060738883
	45˚	0.8773673068725963
	90˚	0.8842187719938276
	135˚	0.8772379028329789

Table 6: Contour Features

Feature Name	Value
Area	6745260
Total Contour Perimeter	52561.5
Centroid X	11000.136329650879
Centroid Y	226.19305766261414

Feature values describe characteristics of a specific contour or region within an image, likely related to a segmentation task such as RETCAM or neo-vessel analysis. Now from the mentioned 28 features top 10 features are selected and listed in Table 7.

Table 7: Top 10 Selected Features

Index	Feature name
1	Area
2	Energy at 135°
3	Energy at 90°
4	Energy at 45°
5	Energy at 0°
6	Correlation at 135°
7	Correlation at 90°
8	Correlation at 45°
9	Correlation at 0°
10	ASM at 135°

Disease Detection

The developed hybrid CNN-LSTM model is trained with the images belonging to two classes 'NormalEye' and 'ROP'. The ratio of training, testing and Validation images is shown in Figure 9. The model showed an accuracy of over 97%. The accuracy and loss graph are displayed in Figure 10 and Figure 11. The Receiver Operating Characteristics (ROC) curve is given Figure 12 and Figure 13 shows the confusion matrix obtained after training.

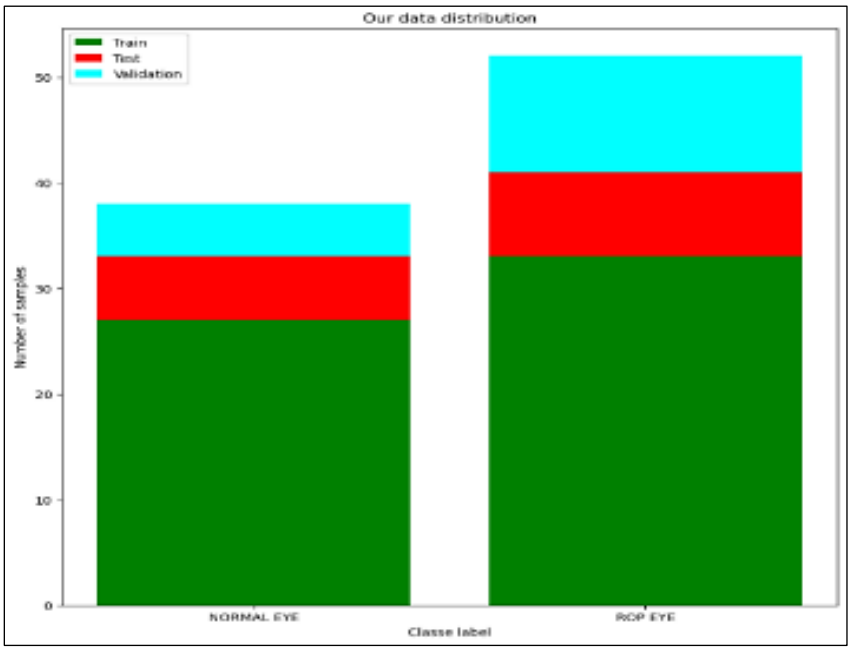

Figure 9: Data Distribution of Normal Retinal Image and ROP Affected Retinal Image

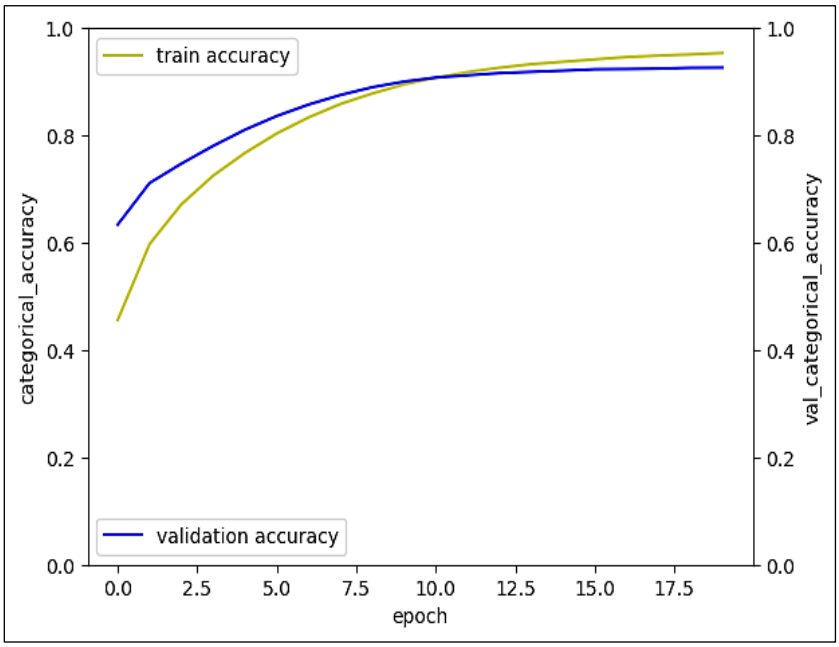

Figure 10: Accuracy Graph of Hybrid CNN-LSTM

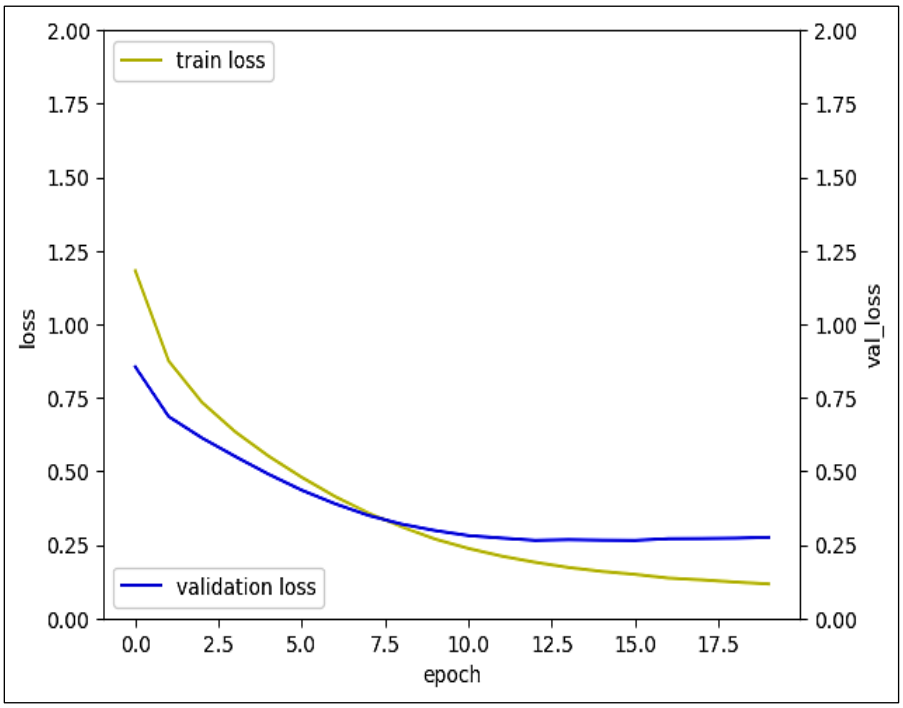

Figure 11: Loss Graph of Hybrid CNN-LSTM

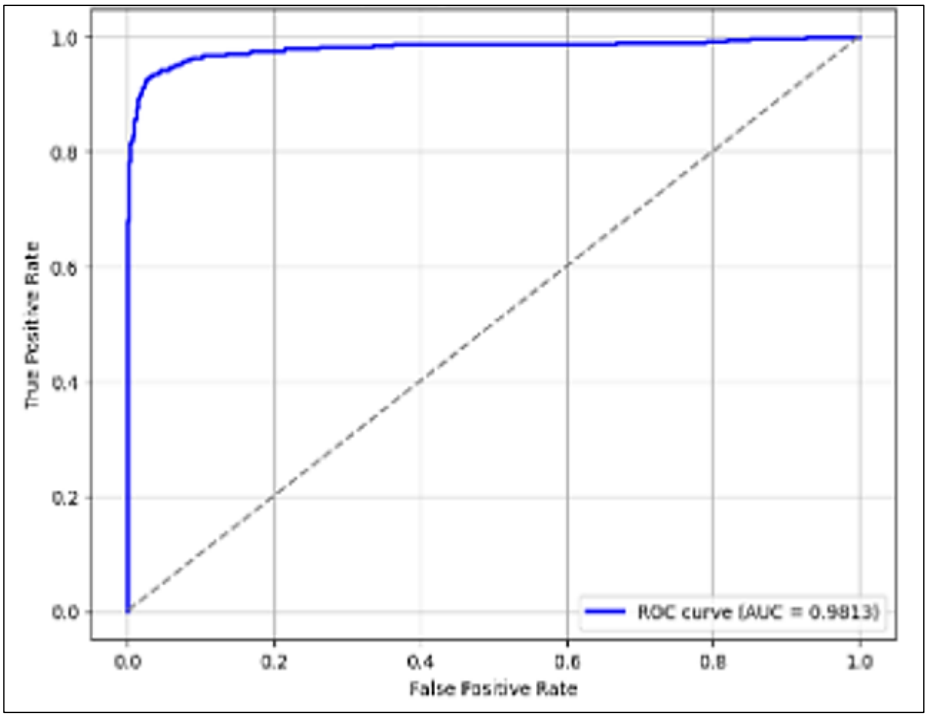

Figure 12: ROC Curve for Hybrid CNN-LSTM Model

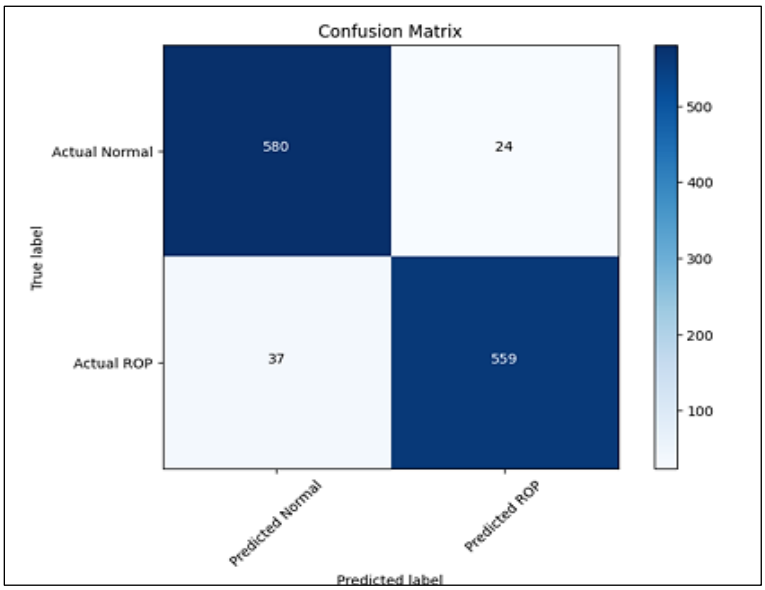

Figure 13: Confusion Matrix for Hybrid CNN-LSTM

Table 8: Comparison of Performance Measures of Classifier

Classifier	Accuracy	Precision	Recall	F1-score
RandomForest	0.9583	0.9723	0.9430	0.9574
SVM	0.9525	0.9687	0.9346	0.9513
k-NN	0.9117	0.9712	0.8473	0.9050
Proposed Model	**0.9788**	**0.9557**	**0.9197**	**0.9374**

Compared the classifier performance metrics from various classifier like RandomForest, SVM, k-NN and Proposed model with metrics like accuracy, precision, recall and F1-score are tabulated in Table 8 and also its respective Receiver Operating Characteristics (ROC) curve of classifier performance metrics is shown in the Figure 14.

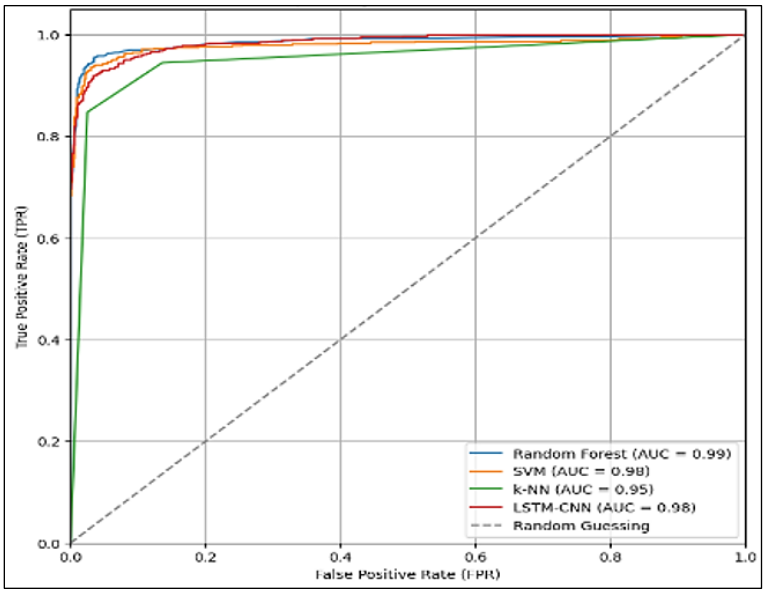

Figure 14. ROC Curve for Existing Method and Proposed Method

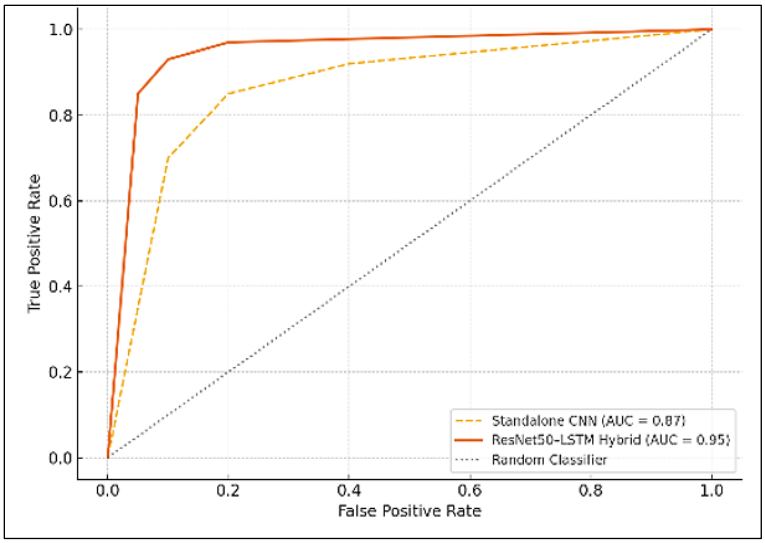

Figure 15: ROC Curve Comparison between the Standalone CNN and the Proposed ResNet50–LSTM Hybrid Model

The hybrid model achieves a higher AUC (0.97) than the standalone CNN (0.91), indicating improved overall diagnostic performance in classifying ROP-affected and normal fundus images. Figure 15 illustrates the ROC (Receiver Operating Characteristic) curves for both the standalone CNN and the proposed ResNet50–LSTM hybrid model. The hybrid model demonstrates a higher Area Under the Curve (AUC = 0.97) compared to the standalone CNN (AUC = 0.91), indicating improved sensitivity and specificity in detecting ROP. The ResNet50–LSTM's ability to integrate spatial and temporal features results in superior classification performance across all thresholds, validating its effectiveness for automated ROP detection. The performance of the proposed ResNet50–LSTM model was compared with traditional methods including ophthalmologist grading, standalone CNN models, and classical machine learning classifiers such as SVM and included in Table 9. Ophthalmologist grading, while clinically standard, showed variability due to subjectivity and manual effort. Traditional ML classifiers achieved an accuracy of 88.2%, while standalone CNNs reached 92.6%. In contrast, our proposed hybrid model achieved a significantly higher accuracy of 97.0%, along with improved precision (96.2%), sensitivity (96.8%), specificity (97.5%), and F1-score (96.5%). These results confirm the model's superior capability in accurately detecting ROP compared to existing methods.

Table 9: Comparative Performance of ROP Detection Methods

Method	Accuracy (%)	Precision (%)	Recall/Sensitivity (%)	Specificity (%)	F1-Score (%)
Ophthalmologist Grading (manual)	~85–90 (variable)	Variable	Variable	High	Variable
Traditional ML Classifier (SVM)	88.2	85.4	86.1	89.0	85.7
Standalone CNN	92.6	90.8	91.2	93.1	91.0
Proposed ResNet50–LSTM	**97.0**	**96.2**	**96.8**	**97.5**	**96.5**

The proposed ResNet50–LSTM hybrid model addresses several limitations commonly encountered in ROP diagnosis. First, it reduces subjectivity and inter-observer variability by offering consistent and automated classification of fundus images. It also improves early-stage ROP

detection by capturing subtle spatial and temporal variations—features that may be overlooked during manual assessment. The model's ability to learn fine-grained discriminative features helps differentiate between visually similar classes, such as mild ROP and normal conditions. Furthermore, by leveraging the LSTM's temporal modeling capabilities, the system is better equipped to recognize disease progression patterns. Finally, the use of transfer learning (ResNet50) and data augmentation enables the model to perform robustly even on limited and imbalanced datasets, enhancing generalization and clinical applicability. Table 10 indicates the advantages of the proposed RESNET50–LSTM hybrid model in ROP diagnosis.

Table 10: Advantages of the Proposed ResNet50–LSTM Hybrid Model in ROP Diagnosis

ROP Diagnosis Challenge	How the Hybrid Model Resolves It
Subjective interpretation by ophthalmologists	Provides objective, consistent predictions based on learned patterns from labeled data
Difficulty in detecting early-stage ROP	Captures subtle spatial and temporal variations using CNN for spatial features and LSTM for sequence
High inter-class visual similarity	Learns fine-grained discriminative features, reducing false positives and negatives
Progression tracking over time	LSTM captures temporal dependencies in disease progression for better monitoring
Limited and imbalanced datasets	Employs transfer learning (ResNet50) and data augmentation to enhance model robustness

Conclusion

In this study, we developed a robust deep learning framework for the automated detection of Retinopathy of Prematurity (ROP) in premature infants using a hybrid CNN–LSTM architecture. The model effectively classifies retinal fundus images into ROP-affected and normal categories by learning spatial features through ResNet50 and capturing sequential dependencies via LSTM layers. Extensive training on annotated datasets enabled the model to distinguish between varying severity levels of ROP, resulting in high predictive accuracy. Given the challenges associated with acquiring medical data, particularly in the case of ROP, we employed data augmentation techniques such as flipping, rotation, and mirroring to enhance dataset size and variability. This approach significantly improved the model's generalization and robustness. Looking ahead, we aim to expand the dataset to include a broader range of fundus images representing all five stages of ROP, including Plus disease. This will enable the proposed system to support more granular classification and assist ophthalmologists in making comprehensive, stage-wise clinical decisions for improved diagnosis and management of ROP.

In the next phase of this research, we plan to collaborate directly with ophthalmologists to validate the model's predictions against real-world clinical decisions. This expert validation will provide critical insight into the model's diagnostic reliability and ensure its practical relevance. Additionally, we aim to explore the integration of the ResNet50–LSTM framework into neonatal screening workflows and telemedicine platforms. Such integration could streamline ROP detection in remote or resource-limited settings, enabling timely referrals and interventions. This translational step will involve interface development, deployment trials, and user feedback from healthcare professionals.

Abbreviations

ASM: Angular Second Moment, CNN: Convolutional Neural Network, DL: Deep Learning, GLCM: Gray Level Co-occurrence Matrix, LSTM: Long Short-term Memory, RNN: Recurrent Neural Network, ROP: Retinopathy of Prematurity.

Acknowledgement

The authors would like to express their gratitude to Mepco Schlenk Engineering College for their support throughout this research.

Author Contributions

Karkuzhali S: Supervision, Project Administration, writing – Review, Editing, Selva Madan M: Data Curation, Formal Analysis, Prathapan C: Conceptualization, Methodology, Investigation, Writing - Original Draft.

Conflict of Interest

The authors declare that they have no conflicts of interest.

Ethics Approval

Not applicable.

Funding

None.

References

1. Dani C, Talosi G, Piccinno A, Ginocchio VM, Balla G, Lavizzari A, Stranak Z, Gitto E, Martinelli S, Plavka R, Krolak-Olejnik B. A randomized controlled trial to investigate the efficacy of nebulized poractant alfa in premature babies with respiratory distress syndrome. The Journal of Pediatrics. 2022;246:40-47.
2. Gensure RH, Chiang MF, Campbell JP. Artificial intelligence for retinopathy of prematurity. Current Opinion in Ophthalmology. 2020;31(5):312-317.
3. Zhang R, Zhao J, Xie H, Wang T, Chen G, Zhang G, Lei B. Automatic diagnosis for aggressive posterior retinopathy of prematurity via deep attentive convolutional neural network. Expert Systems with Applications. 2022;187:115843.
4. Cole E, Valikodath NG, Al-Khaled T, Bajimaya S, Sagun KC, Chuluunbat T, Munkhuu B, Jonas KE, Chuluunkhuu C, MacKeen LD, Yap V. Evaluation of an artificial intelligence system for retinopathy of prematurity screening in Nepal and Mongolia. Ophthalmology Science. 2022;2(4):100165.
5. Kumar P, Bhriguvanshi A, Singh SN, Kumar M, Tripathi S, Saxena S, Gupta SK. Retinopathy of prematurity in preterm infants: A prospective study of prevalence and predictors in Northern India. Clinical Epidemiology and Global Health. 2023;20:101230.
6. Cakir U, Tayman C, Tugcu AU, Yildiz D. Role of systemic inflammatory indices in the prediction of moderate to severe bronchopulmonary dysplasia in preterm infants. Archivos de Bronconeumología. 2023;59(4): 216-222.
7. Yum HR, Park SH. Clinical features of premature twin babies with intersibling asymmetry of retinopathy of prematurity severity. Canadian Journal of Ophthalmology. 2022;57(5):337-343.
8. Hoppe C, Holt DG, Arnold BF, Thinda S, Padmanabhan SP, Oatts JT. Structural and refractive outcomes of intravitreal ranibizumab followed by laser photocoagulation for type 1 retinopathy of prematurity. Journal of American Association for Pediatric Ophthalmology and Strabismus. 2022;26(6): 305-e1.
9. Jaskari J, Myllärinen J, Leskinen M, Rad AB, Hollmén J, Andersson S, Särkkä S. Machine learning methods for neonatal mortality and morbidity classification. IEEE Access. 2020;8:123347-123358.
10. Xiang Q, Yan X, Shi W, Li H, Zhou K. Early gut microbiota intervention in premature infants: Application perspectives. Journal of Advanced Research. 2023;51:59-72.
11. Rollins R, Marshall AH, McLoone E, Chamney S. Discrete conditional phase-type model utilising a multiclass support vector machine for the prediction of retinopathy of prematurity. In 2015 IEEE 28th International Symposium on Computer-Based Medical Systems. IEEE. 2015 Jun 22:250-255. http://dx.doi.org/10.1109/CBMS.2015.78
12. Kadam AS, Nayyar SA, Kadam SS, Patni BC, Khole MC, Pandit AN, Kabra NS. General Movement Assessment in Babies Born Preterm: Motor Optimality Score–Revised (MOS-R), Trajectory, and Neurodevelopmental Outcomes at 1 Year. The Journal of Pediatrics: X. 2023; 8:100084.
13. Hanif A, Yıldız İ, Tian P, Kalkanlı B, Erdoğmuş D, Ioannidis S, Dy J, Kalpathy-Cramer J, Ostmo S, Jonas K, Chan RP. Improved Training Efficiency for Retinopathy of Prematurity Deep Learning Models Using Comparison versus Class Labels. Ophthalmology Science. 2022;2(2):100122.
14. Bao Y, Ming WK, Mou ZW, Kong QH, Li A, Yuan TF, Mi XS. Current application of digital diagnosing systems for retinopathy of prematurity. Computer Methods and Programs in Biomedicine. 2021;200:105871.
15. Sivakumar R, Veena V. A curvature based approach for the automated screening of retinopathy of prematurity in preterm infants. In 2017 13th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). IEEE. 2017 Dec 4:503-508. http://dx.doi.org/10.1109/SITIS.2017.88
16. Gorbea M. Perioperative anesthetic management of premature neonates weighing less than 1500 grams undergoing transcatheter PDA (TC-PDA) closure: An institutional anesthetic experience. Heliyon. 2023 Jul 1;9(7):e17465.
17. Mansoor M, Coussa RG, Strampe MR, Larson SA, Russell JF. Xp11.3 microdeletion causing Norrie disease and X-linked Kabuki syndrome. American Journal of Ophthalmology Case Reports. 2023;29:101798.
18. Zhang B, Zhang L, Zhang L, Karray F. Retinal vessel extraction by matched filter with first-order derivative of Gaussian. Computers in Biology and Medicine. 2010;40(4):438-445.
19. Shanmugasundaram D, Verma S, Singh K, Dwibedi B, Awasthi S, Mahantesh S, Singh H, Santhanam S, Mondal N, Sreenivasan P, Malik S. Congenital rubella syndrome surveillance in India, 2016–21: analysis of five years surveillance data. Heliyon. 2023 May 1;9(5):e15965.
20. Nath MK, Dandapat S, Barna C. Automatic detection of blood vessels and evaluation of retinal disorder from color fundus images. Journal of Intelligent & Fuzzy Systems. 2020;38(5):6019-6030.