iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IV **Month of publication:** April 2025

DOI: https://doi.org/10.22214/ijraset.2025.70060

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Improving Diabetic Retinopathy Detection by using Image Processing and Machine Learning

H. Rajeswary[1], K. Murugan[2], B. Mohammed Umar[3], C. Shanmuga Prasath[4]
[1]*Assistant Professor,* [2]*Head of the Department,* [3,4] *UG Scholars, Department of Electronics and Communication Engineering, RAAK College of Engineering and Technology, Puducherry, India.*

Abstract: ***The rapid development and proliferation of medical imaging technologies is revolutionizing medicine. Medical imaging allows scientists and physicians to glean potentially life-saving information by peering noninvasively into the human body. Diabetic retinopathy (DR) is an irreversible fundus retinopathy. A deep learning-based auto-mated DR diagnosis system can save diagnostic time. This paper reviews and analyses state-of-the-art deep learning methods in supervised, self-supervised and Vision Transformer setups, proposing retinal fundus image classification and detection. For instance, referable, non-referable and proliferative classifications of Diabetic Retinopathy are reviewed and summarized. Moreover, this paper discusses the available retinal fundus datasets for Diabetic Retinopathy that are used for tasks such as detection, classification and segmentation. The paper also assesses research gaps in the area of DR detection/classification and addresses various challenges that need further study and investigation.***

Keywords: Transformer with Multiple Instance Learning (TMIL), Convolution Neural Network (CNN), Global Instance Computing Block (GICB), Vision Transformer (VT).

I. INTRODUCTION

The human eye is a light-sensitive organ that enables us to perceive our surroundings. It functions similarly to a traditional camera, with the image forming on the retina instead of on photographic film. The cornea and crystalline lens of the eye serve the same purpose as a camera lens by focusing incoming light. The iris acts like a camera's diaphragm, regulating the amount of light entering the eye by adjusting the pupil size. Light passes through the cornea, pupil, and lens before reaching the retina at the back of the eye. The retina contains specialized light-sensitive cells called photoreceptors rods and cones which convert light into electrical signals. These signals are sent to the brain via the optic nerve, where they are processed into the visual images we see. The macula, located in the centre of the retina, plays a vital role in sharp central vision. Within it lies the fovea, a region densely packed with cone cells responsible for high-resolution colour vision.

The fovea contains no rod cells, which are more sensitive to brightness and are essential for low-light vision. Cone cells are of three types L, M, and S each sensitive to different wavelengths (long, middle, and short), enabling us to see a broad spectrum of colours. The optic disc is the region where the optic nerve and blood vessels enter the eye. Since it lacks any photoreceptors, it cannot detect light and is referred to as the blind spot. The retina itself is a thin, layered tissue lining the inside of the eye and functions as an extension of the brain, converting light signals into neural signals for interpretation.

Since the retina must remain optically clear to function, the eye's structures are transparent, allowing for non-invasive imaging of the retina. This unique accessibility makes it possible to study brain and circulatory health through the eye. Because the retina is highly metabolically active and has a dual blood supply, it allows for direct observation of blood circulation.

Many eye diseases, such as macular degeneration and glaucoma the leading and third most common causes of blindness manifest in the retina. In addition to eye-specific disorders, systemic diseases can also affect the retina. Diabetic retinopathy, linked to diabetes, is the second most common cause of blindness. Other systemic conditions like hypertension and multiple sclerosis also leave signs in the retina. Imaging the retina not only helps diagnose and monitor eye diseases but also provides insight into broader health conditions. As a result, retinal analysis has become an important tool for early detection and management of both ocular and systemic diseases.

II. DIABETIC RETINOPATHY

Diabetic retinopathy (DR) is a diabetes complication and a major cause of blindness. Hyperglycemia damages retinal vessels, leading to ischemia (causing proliferative DR) and blood-retinal barrier breakdown (causing diabetic macular edema, DME). DME, common in Type-2 diabetes, occurs due to fluid leakage, causing retinal thickening and visual loss. Advanced DME involves cyst formation, tissue damage, and possible serous detachment.

Managing diabetes through controlled blood sugar, lifestyle changes, and anti-diabetic drugs is essential. Treatments like laser photocoagulation, anti-vascular growth factors, and steroids can prevent blindness in DR patients. Regular screening enables early diagnosis and effective intervention.

III. MEDICAL IMAGE PROCESSING

The influence and impact of digital images on modern society is tremendous, and image processing is now a critical component in science and technology. The rapid progress in computerized medical image reconstruction, and the associated developments in analysis methods and computer-aided diagnosis, has propelled medical imaging into one of the most important sub- fields in scientific imaging. Some of the widely used applications involving use of medical imaging are X-ray, CT scan, MRI, ultrasound, retinal image processing etc. X-ray Roentgen's remarkable discovery precipitated one of the most important medical advancements in human history. X-rays are a type of electromagnetic radiation, just like visible light.

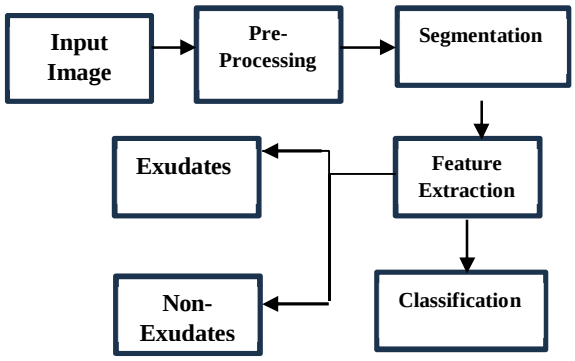


Fig 3.1 Block Diagram of Medical Image Processing

IV. PROPOSED WORK

To develop and implement an advanced, Transformer-based deep learning system for the efficient diagnosis and classification of Diabetic Retinopathy (DR) using high-resolution retinal fundus images. This objective aims to:

1) Overcome the challenges of limited performance on small-scale retinal datasets by utilizing pre-trained weights and a Multiple Instance Learning (MIL) approach.
2) Segment high-resolution retinal images into 448 × 448 patches to preserve critical information and improve the accuracy of DR classification.
3) Leverage Vision Transformer (ViT) models for feature extraction and classification, focusing on referable, non-referable, and proliferative stages of DR.
4) Evaluate and improve upon state-of-the-art supervised and self-supervised learning techniques for DR detection, classification, and segmentation.

A. HBA-U-NET Architecture

The integration of HBA blocks into the U- Net without significantly increasing computational complexity led to the creation of the HBA-U-Net architecture, It consists of:

1) Encoders: These utilize ImageNet pre-trained ResNet-50 blocks to obtain feature maps at different spatial resolutions.
2) Modified U-Net Structure: This structure, equipped with HBA blocks on skip connections processes feature maps alongside the original image to produce the final fovea and OD segmentation mask.
3) Incorporating HBA Blocks: The network has been re-designed to incorporate multiple HBA blocks, creating local bottleneck structures in each skip-connection pair, allowing for different spatial resolutions.
4) The Hierarchical Block Attention (HBA) and HBA-U-Net architecture significantly advance attention mechanisms, particularly in image segmentation. By focusing on individual pixels, spatial relationships, and channel-wise attention, this innovative model presents a refined way of processing images without drastically increasing computational requirements.

B. Working Mechanism of Novel Hybrid ISVM-RBF Model

The Novel Hybrid ISVM-RBF model is a flexible approach that combines an Improved Support Vector Machine with Radial Basis Function (ISVM-RBF), K-Nearest Neighbour (K-NN), and Decision Tree classifiers. This section provides an overview of the hybrid system's fundamental concepts and operational mechanisms in relation to regression and classification tasks. The ISVM-RBF technique may be classified as a non-parametric approach, distinguishing it from traditional statistical parametric approaches. It improves the efficiency and accuracy of change detection, especially when dealing with many data samples.

Kernel Trick: The SVM-RBF transforms variables into a high-dimensional feature space using pre-selected nonlinear mapping functions, creating a more effective classification hyperplane. This nonlinear transformation helps in both linear and nonlinear separating conditions. Radial basis functions (Gaussian variants) are the primary kernel functions used for Mathematical Representation

The ISVM-RBF can be represented by the following equation

$$\text{ISVM-RBF}=\forall\omega(a,a_i) = \exp(-\frac{1}{\sigma^2}||a,a_i||^2) \times \lambda$$

$$\forall\omega(a,P) = (\varphi(a).\varphi(p))$$

$$\forall\omega(a,p) - \{\varphi(a).\varphi(p)\} = \{(a).\varphi(p) - \forall\omega(a,p)\}$$

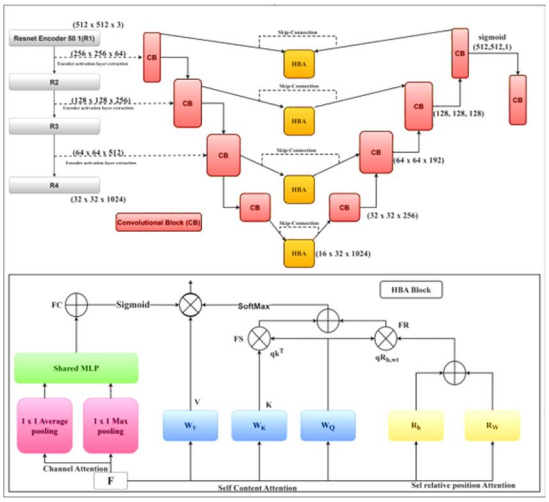


Fig 4.1- Block Diagram of Simulation Process

C. Feature Extraction And Reduction Using CNN-SVD

Identifying and classifying various stages of Diabetic Retinopathy (DR) requires a robust approach that combines feature extraction with feature reduction. Initially, a Convolutional Neural Network (CNN) is used to extract a wide range of Fundamental Image (FI) features from retinal images. This CNN model targets features that help distinguish the different stages of DR.

The CNN includes layers such as batch normalization, max-pooling, and dropout. Batch normalization speeds up training and improves model performance by stabilizing layer inputs. Max-pooling reduces data complexity by focusing on the most important features, while dropout prevents overfitting by randomly deactivating neurons during training. The Adam optimizer is employed for its efficiency and adaptability with large datasets. At the final dense layer of the CNN, 256 attributes are extracted from each image to support classification. Once these features are extracted, the model applies a feature reduction technique based on Fast Fourier Transform (FFT) and Singular Value Decomposition (SVD).This involves representing the feature matrix of size m×nm \times nm×n as the product of three matrices: A=STU. Each matrix has a specific role: the matrix captures key image structures like edges and textures, the SSS matrix quantifies the importance of each spatial feature, and the matrix holds the singular values. SVD is then used to further decompose this structure. By selecting the top singular values and the corresponding vectors, the model retains only the most relevant data, significantly reducing dimensionality while preserving key information. This decomposition relies on mathematical principles involving unitary and diagonal matrices to maintain the integrity of important image features. Combining CNN-based feature extraction with SVD-driven feature reduction results in a powerful, efficient model.

This approach simplifies data complexity while preserving critical diagnostic features, making it highly effective for classifying DR stages. It offers a well-balanced, accurate, and computationally efficient method for aiding the diagnosis and management of DR.

$$stu = \begin{bmatrix} s_1 & s_2 \\ s_3 & s_4 \end{bmatrix}\begin{bmatrix} t_1 & 0 \\ 0 & t_2 \end{bmatrix}\begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix} = \begin{bmatrix} s_1t_1 & s_2t_2 \\ s_3t_1 & s_4t_2 \end{bmatrix}\begin{bmatrix} u_1 & u_2 \\ u_3 & u_4 \end{bmatrix}$$

$$= \begin{bmatrix} s_1t_1u_1 + s_2t_2u_3 & s_1t_1u_2 + s_2t_2u_4 \\ s_3t_1u_1 + s_4t_2u_3 & s_3t_1u_2 + s_4t_2u_4 \end{bmatrix} = t_1\begin{bmatrix} s_1 \\ s_3 \end{bmatrix}\begin{bmatrix} u_1 & u_2 \end{bmatrix} + t_2\begin{bmatrix} s_2 \\ s_4 \end{bmatrix}\begin{bmatrix} u_3 & u_4 \end{bmatrix}$$

Hence, PQR* can be written as:

$$\begin{bmatrix} | & & | \\ S_1 & S_2 \cdots & S_m \\ | & & | \end{bmatrix}\begin{bmatrix} T_1 & 0 & 0 \\ 0 & T_2 & 0 \\ | & | & T_m \end{bmatrix}\begin{bmatrix} - & U_1 & - \\ - & U_2 & - \\ & \vdots & \\ - & U_n & - \end{bmatrix}$$

$$= T_1\begin{bmatrix} | \\ S_1 \\ | \end{bmatrix}\begin{bmatrix} - & U_1 & - \end{bmatrix} + T_2\begin{bmatrix} | \\ S_2 \\ | \end{bmatrix}\begin{bmatrix} - & U_2 & - \end{bmatrix} + \cdots + T_m\begin{bmatrix} | \\ S_m \\ | \end{bmatrix}\begin{bmatrix} - & U_n & - \end{bmatrix}$$

The main aim of the proposed approach is to effectually classify the grade of diabetic retinopathy using retinal images. DR is based on chronic complications of retinal blood vessels. Early-stage prediction and diagnosis prevent vision loss. Therefore, in this work, we focus on automatic diabetic retinopathy detection. This process is helpful for doctors to give the proper medicine to patients. The two stages of the suggested methodology are pre-processing and classification. The images are initially fetched from the database. The pre-processing stage then receives the collected images. The acquired images are upgraded utilizing the following steps during the pre-processing stage: noise removal using a Gaussian filter, image enhancement using CLAHE, blur removal using an un-sharp filter, and image resizing using cropping. The classifier receives the images right away following pre-processing. In this study, the two-way cascade CNN is introduced for classification. The input image's local and global features are retrieved in this approach, and both feature maps are integrated. The classifier generates four separate classes, including none, mild, moderate, and severe diabetic retinopathy.

D. Two Layer Neural Network

The Indian Diabetic Retinopathy Imaging Dataset (IDRiD) and the Kaggle Diabetic Retinopathy Detection Training Dataset (KDRDTD), both of which comprise n pictures, are used to collect the retinal images. The images taken by IDRiD have a field of view of 50° and a resolution of 4288 × 2848 pixels; they are saved in jpg format. The final dataset consists of 516 images that have been divided into 3 categories for diabetic macular edema and 5 categories for diabetic retinopathy using stringent criteria based on global clinical criteria. It consists of professional comment.

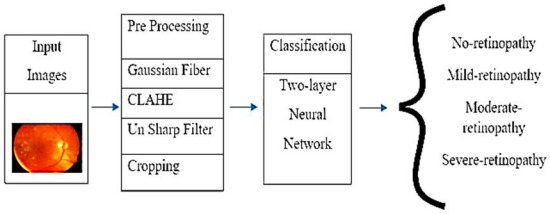


Fig 4.2-Diabetic retinopathy classification process

On typical DR lesions as well as typical retinal elements. This additionally offers DR and DME disease severity levels for every image within the database. The KDRDTD dataset is a massive collection of high-quality retina images captured during various imaging settings. In each subject, left and right fields are given. In addition to the left or right eye, images can be recognized by a subject ID (for instance, 1_left.jpeg indicates patient number 1's left eye). Every image was rated by a doctor for the presence of diabetic retinopathy on a scale of 0 to 4, with No DR, Severe, Mild and Moderate. Some of the sample images are presented.

E. Diabetic Retinopathy Grading using two Layer Neural Network CASCADE CNN

A two-path CNN classifier is proposed for diabetic retinopathy (DR) detection, leveraging both local and global image features. The shallow path uses large kernels for global features, while the deep path employs small kernels for local details. Combined features are processed through convolutional and fully connected layers for classification. This approach improves DR stage prediction accuracy. Various studies using CNN, transfer learning, and deep learning methods show high performance in DR detection. Techniques like data augmentation, ensemble learning, and AC-Dense Net enhance accuracy and robustness. The best models achieved up to 99.36% accuracy, aiding early diagnosis and reducing human error.

F. Performance Evaluation

The performance of our model is evaluated using various metrics. These evaluation metrics are used to gauge the efficiency and accuracy of a RSG-Net predictions. When evaluating both the training and testing sets in this study, each data instance can fall into one of four categories: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). A True Positive denotes the model correctly predicting a positive

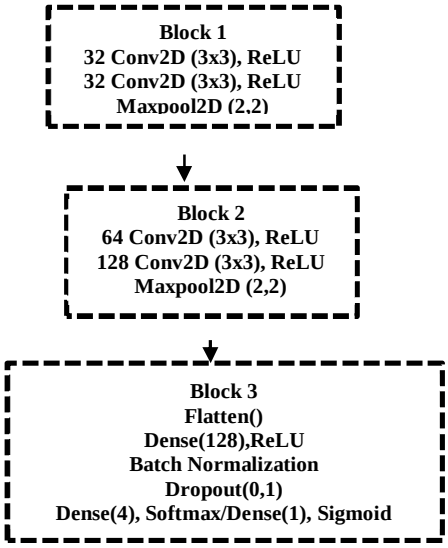


Fig 4.3- Detail on the layers used for RSG-Net

V. SOFTWARE AND HARDWARE REQUIREMENTS

A. Software Requirements

Operating System: Windows 7/ 8/ 10
Software Package: MATLAB 2014a
Software: Matlab8.3

B. Hardware Requirements

Processor: Dual Core2.0
Hard Disk Space: Min 10Gb (Package and Library Extraction)
Physical Memory: 2-4 Gb
Input: Standard Pointing device and Keying device
Output: Standard display

C. Platform Configuration

The platform configuration is tabulated below

Table 5.1- Platform Configuration

Variable	Configuration/ Description
Observation Frequency	50Hz
Testing Interval Time (s)	30
Samples	5
Variations	0.04
Max Time (s)	80

VI. VIRTUAL OUTPUT

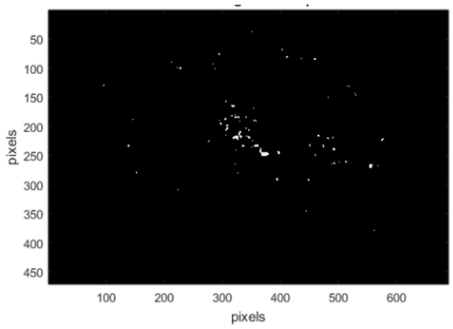


Fig 6.1- Final Segmentation for affected region

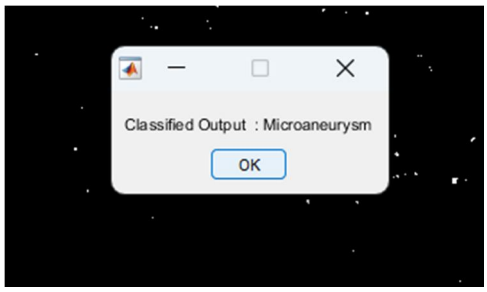


Fig 6.2- Displaying the disease type 1

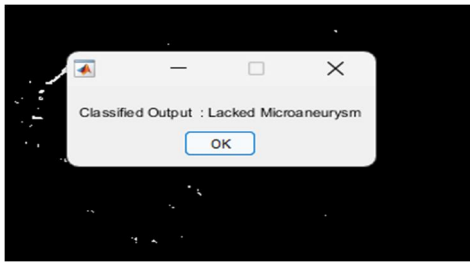


Fig 6.3- Displaying the disease type 2

This project is designed to detect diabetic retinopathy by identifying microaneurysms in retinal fundus images. To achieve accurate detection, the images undergo pre-processing to enhance quality and isolate key retinal features. The resulting "Final segmented Op" image represents the segmented output, where small bright spots may indicate the presence of microaneurysms. The system provides two possible classification results: “Lacked Microaneurysm,” indicating no signs of diabetic retinopathy, and “Microaneurysm,” confirming the presence of early-stage retinopathy. The output is presented visually through a segmented image with affected regions highlighted in white against a black background, and textually via a pop-up dialog box displaying the classification label.

```
-----------------------------------------
            Actual Classes
------------1--------2-----
    TP |   27.00 |   21.00
    FP |   2.00 |   0.00
    FN |   0.00 |   2.00
    TN |   21.00 |   27.00
Preci. |   0.93 |   1.00
Sensi. |   1.00 |   0.91
Speci. |   0.91 |   1.00
-----------------------------------------
Model Accuracy is 0.96
-----------------------------------------
```

Fig 6.4 - Accuracy Outcome

The model achieves an overall accuracy of 96%, demonstrating its robustness in diabetic classification

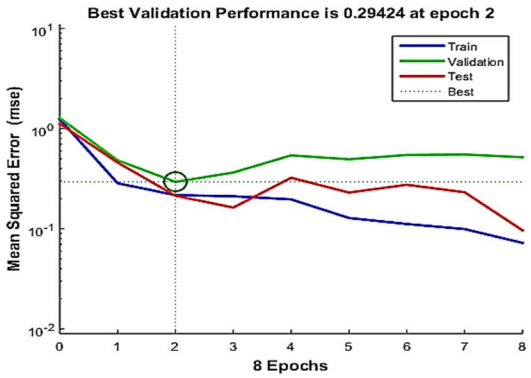


Fig 6.5 – Best Validation performance

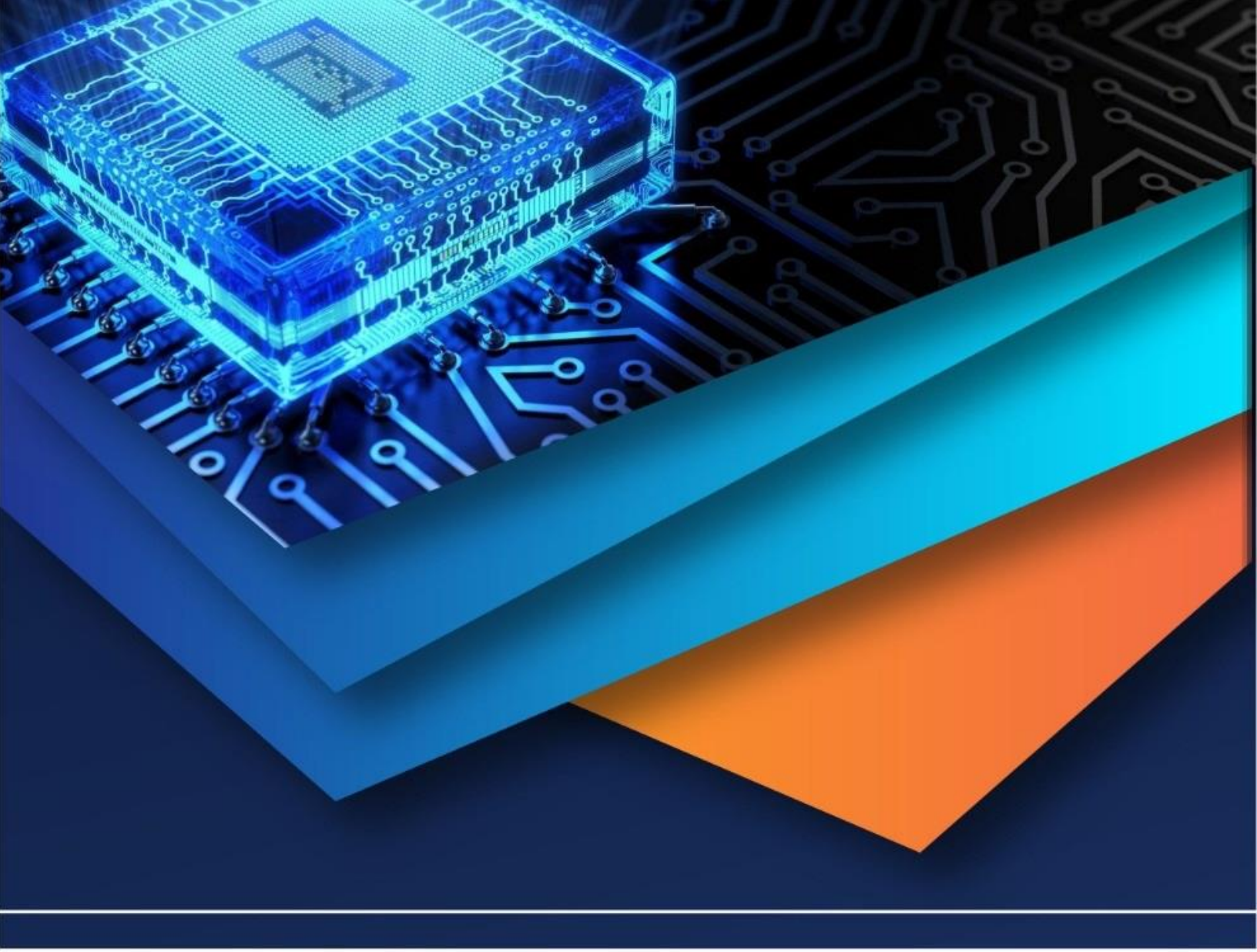
1) The blue line represents the Training dataset, showing the model's gradual adaptation to the data as epochs progress.
2) The green line represents the Validation dataset, assessing the model's ability to generalize to unseen data.
3) The red line signifies the Test dataset, used for final evaluation post-training and validation phases.

VII. CONCLUSION AND FUTURE SCOPE

This project presents an efficient method for detecting blood vessels and hard exudates in retinal fundus images, aiding in the early diagnosis of diabetic retinopathy. Image segmentation techniques were applied effectively, with performance influenced by factors like intensity and texture. The system demonstrated reduced human error and suitability for remote diagnostics with low computational cost. It serves as a valuable screening tool for early DR detection by identifying key clinical features such as vessels, exudates, and optic disc. Future work should focus on precise exudate extraction using a hybrid of supervised and unsupervised methods. Combining CNNs with clustering techniques can enhance accuracy, reduce bias, and improve detection scalability. This approach could significantly advance diabetic retinopathy screening and broader medical imaging applications.

REFERENCES

[1] V.Kumari and N.Suriyanarayanan, "Feature Extraction for Early Detection of Diabetic Retinopathy," In International Conference on Recent trends in Information, Telecommunication and Computing, 2010, Pp. 359-361.
[2] S.Ravishankar, "Automated Feature Extraction for Early Detection of Diabetic Retinopathy in Fundus Images" In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference On 2009, Pp. 210-217.
[3] Neeraj Sharma and Lalit M.Aggarwal, “Automated Medical Image Segmentation Techniques” Journal of Medical Physics, 2010 Jan-Mar
[4] K.V.Kauppi T, Lensu L, Sorri I, Raninen A, "Diaretdb1: Diabetic Retinopathy Database and Evaluation Protocol. In: Medical Image Understanding and Analysis (MIUA).” Ed, 2007.
[5] Younis N, Broadbent Dm, Harding Sp, Vora Jp. Incidence Of Sight-Threatening Retinopathy in Type 1 Diabetes in a Systematic Screening Programme. Diabet Med ; 20:758-65.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY