

## Firefly-Based Segmentation and Residual Deep Learning for Multi-Class Diabetic Retinopathy Detection

Prathibha S.<sup>[1]</sup>, Madhusudhan K. N.<sup>[1]</sup>, Agughasi Victor Ikechukwu<sup>[2]</sup>

<sup>[1]</sup> Department of Electronics & Communication Engineering, BMSCE Bengaluru, Affiliated to Visvesvaraya Technological University, Belagavi, 580018, India. prathi.gecr@gmail.com, madhusudhankn.ece@bmsce.ac.in

<sup>[2]</sup> Department of CSE (Artificial Intelligence), Maharaja Institute of Technology Mysore, Affiliated to Visvesvaraya Technological University, Belagavi, 580018, India. victor.agughasi@gmail.com

**Abstract** In recent years, the rise in diabetic retinopathy cases has posed significant challenges to existing computer-aided diagnosis (CAD) solutions. These systems often focus on detecting specific abnormalities, such as microaneurysms, exudates, or hemorrhages, rather than providing a comprehensive diagnostic approach. Moreover, state-of-the-art deep learning-based methods face critical limitations, including a lack of contextual understanding, gradient vanishing/explosion issues, and failure to address class imbalance at the instance level, which impacts multi-class classification accuracy. To overcome these challenges, a novel diabetic retinopathy prediction model is proposed, leveraging firefly heuristic segmentation and residual deep spatio-textural feature learning. Instead of processing entire fundus images, the model applies Firefly heuristic-driven Fuzzy C-Means (FFCM) clustering to segment regions of interest (ROIs) corresponding to microaneurysms, exudates, and hemorrhages. Residual deep-spatio-textural features are then extracted using Gray-Level Co-occurrence Matrix (GLCM), ResNet50, and AlexNet. These complementary features enhance diversity and heterogeneity, which are further processed using random forest learning. The proposed model achieves outstanding performance, with an average accuracy of 99.77%, precision of 99.88%, recall of 99.64%, F-measure of 99.75%, sensitivity of 99.64%, and specificity of 99.86%, surpassing existing approaches. FFCM mitigates the class imbalance problem, ResNet50 addresses gradient challenges, and AlexNet contributes high-dimensional features, ensuring robust and scalable diagnostics. This innovative solution demonstrates exceptional generalizability and runtime efficiency, offering a cost-effective, comprehensive CAD tool for diabetic retinopathy detection.

**Keywords:** Diabetic Retinopathy Detection, Firefly Heuristic Clustering, Residual Deep Learning, Spatio-Textural Feature Extraction, Class Imbalance Mitigation

### 1. Introduction

Diabetic retinopathy (DR) has emerged as a critical global healthcare challenge due to the increasing prevalence of diabetes, driven by rapid population growth and lifestyle changes[1]. Among various complications caused by diabetes, DR remains one of the most severe, often leading to vision impairment and blindness [2]. The primary contributors to diabetes-related health issues include uncontrolled metabolism, poor insulin regulation, unhealthy living conditions, alcohol consumption, and poor dietary habits. The impact of diabetes is not limited to any single organ but manifests systemically, with the retina being one of the most severely affected regions. DR-related complications include color blindness, myopia, hypermetropia, loss of vision, and other retinal malfunctions [3],[4]. Medical science attributes these conditions to the biochemical disruptions caused by diabetes, which damage the retinal structure and functionality over time.

The retina is a thin, spherical, and innermost layer of the eye that plays a crucial role in visual perception by converting light into neural signals. These signals are processed by the brain to form perceptible visual information. Structurally, the retina surrounds the macula, a dark circular area located near the optical nerves. Within the macula lies the fovea centralis, which facilitates detailed vision required for decision-making [4]. Any disruption in the delicate structural and functional dynamics of the retina, whether due to nerve damage or vascular irregularities, can significantly impair vision. Diabetes alters the biochemical and structural properties of the retina, causing abnormalities such as swelling, clogging of blood vessels, and disrupted blood flow. These changes give rise to conditions such as microaneurysms, exudates, and hemorrhages, which are hallmarks of DR progression [5-7]. Literature [8] suggests that approximately 400 million people worldwide suffer from diabetic retinopathy, a number that continues to grow at an alarming pace [2][6]. Consequently, DR has become one of the leading causes of vision impairment and blindness globally [2].

Early diagnosis and treatment of DR are essential to prevent or mitigate its progression and the associated risks of vision loss. However, detecting DR in its early stages remains a formidable challenge due to the lack of noticeable symptoms and quantifiable biomarkers during the initial phases of the disease. Traditionally, DR diagnosis relies on ophthalmologists examining retinal fundus images to identify abnormalities. While manual assessments can provide accurate diagnoses, they are time-intensive and prone to human error, especially given the non-linear and complex structures associated with retinal abnormalities such as microaneurysms, exudates, hemorrhages, and macular detachment [5]. These complexities reduce the scalability of manual methods, making them insufficient to meet the increasing demand for early DR diagnosis. Additionally, manual assessments are ill-suited for large-scale screenings due to their reliance on specialized expertise and the time required to analyze each image.

To address the limitations of manual diagnosis, computer-aided diagnosis (CAD) systems have gained attention as a promising alternative. CAD systems employ advanced computational techniques, including artificial intelligence (AI) and machine learning, to analyze retinal fundus images and automate the detection of DR-related abnormalities [8]. These solutions aim to improve diagnostic accuracy, scalability, and efficiency. However, several challenges remain. The inherent variability in retinal abnormalities—ranging from microaneurysms to exudates and hemorrhages—introduces significant complexity in feature extraction and classification. These abnormalities differ not only in size, shape, and color but also in textural and spatial characteristics, making it difficult for CAD systems to generalize across diverse datasets. Additionally, retinal images often exhibit high non-linearity, and their complex structures can lead to false positives or negatives during analysis.

Another critical challenge in developing CAD systems for DR diagnosis is class imbalance, where certain abnormalities are underrepresented in datasets. This imbalance can bias machine learning models toward more frequently occurring abnormalities, reducing the overall diagnostic accuracy. Moreover, many existing CAD solutions focus on detecting specific abnormalities, such as microaneurysms or exudates, without addressing the need for a comprehensive diagnostic approach that can simultaneously detect multiple DR-related issues. The reliance on conventional image processing techniques, such as wavelet analysis, Gabor filters, and local binary patterns (LBP), further limits the scalability of these systems [8][9]. These methods often depend on predefined thresholds, such as Otsu's method, for region-of-interest (ROI) segmentation and feature extraction. While effective in controlled settings, static thresholding methods struggle to handle the variability and non-linearity present in real-world datasets, leading to suboptimal results and reduced reliability [9][10].

Deep learning (DL) methods, particularly convolutional neural networks (CNNs) and residual networks (ResNet50), have shown significant promise in addressing these challenges. These approaches excel in extracting hierarchical features from complex data, making them well-suited for analyzing retinal fundus images. However, most deep learning models for DR diagnosis focus on specific abnormalities, such as microaneurysms, exudates, or hemorrhages, and do not provide a unified framework for comprehensive DR detection. Furthermore, CNN-based models often suffer from issues such as overfitting, gradient vanishing, and a lack of contextual awareness, which limit their applicability to

---

diverse datasets. The use of entire fundus images as input to convolutional layers can also lead to computational inefficiencies and reduced performance, as irrelevant features may dilute the effectiveness of the extracted features.

To bridge the gap between the clinical motivation and the technical core of this study, we now outline the complete solution architecture. The proposed pipeline begins with Firefly-optimised Fuzzy C-Means (FFCM) segmentation to isolate retinal lesions, continues with a hybrid feature extractor that concatenates handcrafted GLCM texture descriptors with deep residual embeddings from a fine-tuned AlexNet backbone, and culminates in a Random-Forest classifier that outputs both lesion-level maps and image-level DR severity scores (Normal, Exudate, Micro-aneurysm, Haemorrhage). Additionally, traditional FCM clustering does not account for the non-linearity and variability inherent in retinal abnormalities, further limiting its effectiveness. This study proposes a novel CAD solution for DR diagnosis that integrates Firefly heuristic-driven Fuzzy C-Means (FFCM) clustering with hybrid residual deep spatio-textural feature learning. FFCM addresses the limitations of traditional FCM by leveraging heuristic optimization to dynamically refine centroids during clustering. This approach enables accurate segmentation of non-linear and irregular ROIs, such as microaneurysms, exudates, and hemorrhages, without requiring manual seed-point definition or predefining the number of clusters. The segmented ROIs are then processed using a hybrid feature extraction framework that combines Gray-Level Co-occurrence Matrix (GLCM) with deep learning models, including ResNet50 and AlexNet. GLCM captures essential spatio-textural features, such as entropy, energy, homogeneity, and contrast, while ResNet50 and AlexNet provide complementary high-dimensional deep features. This hybrid approach ensures feature diversity and robustness, enhancing the model's ability to classify multiple abnormalities simultaneously.

**The Contributions are:**

- **Comprehensive Multi-Type Diagnosis:** The proposed vision-computing-based CAD solution is robust at detecting and classifying multiple diabetic retinopathy abnormalities (microaneurysms, exudates, and hemorrhages) in a single system. This ensures a cost-effective and scalable diagnostic approach suitable for clinical applications.
- **Enhanced ROI Segmentation with FFCM:** The Firefly heuristic-driven Fuzzy C-Means (FFCM) segmentation method overcomes the limitations of traditional approaches by automating ROI segmentation, addressing data imbalance, and eliminating the need for manual seed-point definitions. This enables effective real-time application without predefining the number of clusters or centroids.
- **Hybrid Feature Extraction for Enhanced Learning:** By combining ROI-specific spatio-textural features from GLCM (e.g., entropy, contrast, skewness) with deep features from ResNet50 and AlexNet, the proposed method creates a robust and diverse feature space. This mitigates challenges like gradient vanishing, gradient explosion, and computational inefficiencies, ensuring superior learning and prediction for multi-class classification.
- **State-of-the-Art Performance:** The proposed model achieved the highest performance metrics on benchmark datasets, with 99.77% accuracy, 99.88% precision, 99.64% recall, 99.75% F-Measure, 99.64% sensitivity, and 99.86% specificity, demonstrating better performance for real-time DR diagnosis.

By addressing the challenges of class imbalance, feature diversity, and computational efficiency, the proposed CAD solution offers a scalable and cost-effective framework for DR diagnosis. It provides a comprehensive diagnostic approach that can detect and classify multiple types of retinal abnormalities in a single analysis, reducing the need for multiple tests and enabling faster decision-making. Simulation results demonstrate that the proposed model achieves state-of-the-art performance across key evaluation

---

metrics, including accuracy, precision, recall, sensitivity, and specificity, making it a promising tool for real-time clinical applications.

The subsequent sections of this manuscript are organized as follows: **Section 2** reviews the related works, highlighting existing vision-based techniques for detecting and classifying DR abnormalities such as exudates, microaneurysms (MA), and hemorrhages. This section emphasizes the strengths and limitations of these methods, particularly in handling non-linear and complex retinal structures. **Section 3** introduces the proposed model in detail, outlining its integration of Firefly heuristic-driven Fuzzy C-Means (FFCM) segmentation and hybrid feature learning. It explains how the model addresses the challenges of data imbalance, computational inefficiencies, and gradient-related issues while improving feature diversity for robust classification. **Section 4** presents the simulation results, demonstrating the model's superior performance on benchmark datasets. This section includes a thorough analysis of the metrics, comparing the proposed approach to existing state-of-the-art methods. Finally, **Section 5** discusses the broader implications of the findings, their relevance to real-world applications, and potential avenues for future research. A comprehensive list of references is provided at the end of the manuscript to support and validate the research exploration.

## 2. Related Work

The early detection and classification of diabetic retinopathy (DR) abnormalities, such as microaneurysms (MA), exudates (EX), and hemorrhages (HEM), have been extensively studied in recent years. Researchers have explored various methods, ranging from traditional image processing techniques to deep learning-driven solutions, for retinal segmentation, abnormality detection, and classification. This section reviews notable contributions in these areas, with an emphasis on their methods, limitations, and advancements.

### i Retinal Blood Vessel Segmentation and Analysis

Segmenting retinal blood vessels is a critical task in DR diagnosis, as it aids in identifying abnormalities and regions of interest (ROIs). Leopold et al. [11] employed pixel back-propagation neural networks for retinal segmentation, achieving promising results. Mahapatra et al. [12] proposed a triplet loss function to enhance the quality of fundus images, wherein the improved image output was fed iteratively to subsequent images. However, this approach suffered from high computational complexity, limiting its real-time applicability. Wang et al. [13] developed a cascade design for detecting retinal blood vessels, which improved accuracy but required significant computational resources. Similarly, a tri-map-based method utilizing local attributes of blood vessels for segmentation was proposed by [14], where hierarchical image matting techniques were employed to refine ROI pixel details. Hossain and Reza [15] introduced a Markov Random Field (MRF)-based approach for blood vessel segmentation, leveraging Bayesian rule for joint distribution estimation and Markov-Gibbs equivalence for clique energy measurement. Bandara and Giragama [16] applied the Tyler Coye method, utilizing Hough-line transformation for vessel segmentation, while Adal et al. [17] utilized a multi-scale blob-ness approach for lesion shape detection in fundus images, followed by support vector machine (SVM)-based classification. Costa et al. [18] introduced adversarial auto-encoders (AAE) to segment retinal blood vessels, which showed robustness in handling complex vessel structures. Deep learning-based approaches have also gained prominence in this domain. Maninis et al. [19] implemented a deep convolutional neural network (DCNN) for retinal vessel and optic disc segmentation. Patwari et al. [20] analyzed various morphological aspects of vessels, such as diameter, thickness, and length, for classification tasks. Lahiri et al. [21] employed ensemble unsupervised learning, using denoised stacked auto-encoders (SDAE) to extract features, which were classified using a Softmax classifier trained with cross-entropy loss. Odstreilik et al. [22] focused on vessel diameter information to construct high-quality datasets, later processed with matched filters for vessel segmentation. Relan et al. [23] combined Gaussian Mixture

---

Models (GMM) with expectation maximization for blood vessel segmentation, although computational inefficiency and accuracy limitations were noted. Calvo et al. [24] used morphological operations for vessel detection, while [25] applied double-ring filters and top-hat transformations for ROI segmentation and classification using linear discriminant analysis (LDA).

## ii Microaneurysm (MA) Detection

Microaneurysms, one of the earliest signs of DR, have been extensively studied for early diagnosis. Chowdhury et al. [33] employed K-Means clustering for initial segmentation, followed by ROI-specific feature learning using random forests for classification. Lam et al. [34] utilized CNNs to extract local features from fundus images for detecting MA, exudates, hemorrhages, and normal structures. Similarly, Orlando et al. [35] integrated CNN-based feature extraction with handcrafted spatio-textural features, combining these with random forests for red lesion detection. Cao et al. [36] introduced principal component analysis (PCA) for feature selection, followed by SVM and random forests for MA classification. Srivastava et al. [37] applied Frangi-based filtration for vessel detection, followed by SVM-based learning for MA and hemorrhage detection. Hatanaka et al. [38] proposed a double-ring filter for detecting MA lesions and suggested PCA-driven feature extraction to improve accuracy when combined with artificial neural networks (ANN). Venkatesan et al. [39] used color-correlogram features for detecting neovascularization and MA, utilizing multiple instance learning for feature extraction and supervised classification. Shah et al. [48] emphasized hybrid lesion-specific feature learning as a viable alternative to standalone CNN-based solutions for DR classification. Morphological methods have also been explored extensively for MA detection. Spencer et al. [49] used vasculature removal to isolate MAs, while Walter et al. [50] employed morphological reconstruction for detecting MAs and exudates. Advanced techniques such as adaptive wavelet transforms [56] and residual U-Nets [61] have further improved the segmentation and detection of MAs. Residual U-Nets were found particularly effective in addressing class imbalance while providing robust feature representations. Kou et al. [61] demonstrated that this method, coupled with ResNet for feature learning, significantly improved MA detection in fundus images.

## iii Exudate (EX) Detection

Exudates, which indicate advanced DR stages, have also been a focus of several studies. Perdomo et al. [62] utilized LeNet CNN for exudate detection, while Fujita et al. [64] employed CNNs with cross-entropy loss for detecting exudates, MAs, and hemorrhages. Feng et al. [65] combined FCNN with residual functions to segment optic discs and exudates. Zheng et al. [66] developed an ensemble method named MUNet for exudate detection, demonstrating robust performance across varied datasets. Wu et al. [67] applied visual features with 2D Gaussian fitting for hemorrhage detection. Splat-based feature learning [69][70] was also introduced to classify hemorrhages, leveraging color and spatial location features. Kande et al. [71] utilized morphological functions and pixel-wise classification for detecting hemorrhages and MAs. For exudate detection, several methods rely on texture analysis and machine learning classifiers. Prenta and Lončarić [72][73] employed DCNNs as feature extractors, combining them with SVM classifiers for detection. Omar et al. [74] utilized local binary patterns (LBP) and texture features, training an ANN for exudate classification. Mahapatra et al. [75] proposed a CNN-based framework augmented with handcrafted features, while Osareh et al. [76] applied Fuzzy C-Means (FCM) for exudate segmentation, followed by ANN-based ROI contour estimation. Sopharak et al. [77] merged FCM with hierarchical SVMs (HSVM) for exudate detection, while mathematical morphology algorithms (MMA) were applied by [78] for robust segmentation. Harangi and Hajdu [79] extended MMA with active contour techniques to minimize false positives. Zabihollahy et al. [81] introduced a

---

modified U-Net for hard-exudate segmentation, while Yan et al. [82] improved segmentation accuracy by combining local and global U-Net decoders. Jaafar et al. [83] proposed a two-step automated segmentation method for exudates, involving top-down segmentation and polar coordinate mechanisms for assessing severity. Casanova et al. [84] employed random forest ensembles for diabetic retinopathy assessment, using MA-based segmentation approaches for graded fundus image analysis.

#### iv Emerging Deep Learning Approaches

The use of deep learning has significantly advanced DR diagnosis. Ting et al. [28] highlighted the potential of deep neural networks (DNNs) for retinal blood vessel segmentation, DR detection, and glaucoma classification. Almotiri et al. [30] emphasized that training deep models on ROI-specific features yielded more reliable performance compared to generic approaches. Singh et al. [31] echoed similar sentiments, recommending ROI-specific wavelet features with hyperparameter tuning for accurate glaucoma detection. Fully convolutional neural networks (FCNNs) have also been widely used for DR-related tasks. Chudzik et al. [60] employed FCNNs with batch normalization layers for MA detection, utilizing Dice coefficient loss for improved segmentation efficacy. Similarly, Zhao et al. [68] introduced the R-sGAN model, leveraging synthetic retinal images for supervised retinal segmentation. Hybrid deep learning models have further demonstrated their utility in DR detection. Yan et al. [82] amalgamated local and global decoders within U-Net architectures to enhance segmentation accuracy. Zheng et al. [66] proposed an ensemble approach with MUNet for detecting exudates, which outperformed standalone deep learning methods. Fujita et al. [64] applied CNNs to detect multiple abnormalities, including exudates, MAs, and hemorrhages, using cross-entropy loss to refine feature learning. Advanced ensemble methods, such as the combination of residual functions and UNets proposed by Feng et al. [65], have shown significant improvements in both segmentation and classification. Despite the advancements in DR diagnosis, several challenges persist. Many approaches still struggle with class imbalance, non-linear features, and high computational demands. ROI-specific feature learning has emerged as a potential solution to these challenges, as highlighted by Almotiri et al. [30] and Singh et al. [31]. However, ensuring generalizability across diverse datasets remains an open problem. Furthermore, while DL methods like CNNs, U-Nets, and FCNNs have shown promise, their reliance on large labeled datasets can be a bottleneck. Techniques such as transfer learning, semi-supervised learning, and synthetic data generation (e.g., R-sGAN [68]) may address these limitations. Summary of relevant studies is presented in Table I.

**Table I:** Recent DR studies, summarising data sources, target tasks, and key contributions and or limitations

Study (Year)	Dataset	Lesion/Task	Method	Acc./Dice	Key Limitation	Ref.
Mateen et al. (2020)	Messidor	Exudates	DenseNet121	94% Acc	Ignores microaneurysms	[8]
Kou et al. (2023)	IDRiD	Pixel segmentation	Residual U-Net	0.89 Dice	GPU-heavy	[9]
Wang et al. (2024)	Private	DR prediction	MB-TCN with TDC	AUC 0.95	Missing validation on public set	[4]
Ahmad et al. (2024)	Private	Vessel tracking	Gaussian Process + ROA	Acc. 95.3%	Not lesion-specific	[2]
Jabbar et al. (2024)	IDRiD	Lesion classification	Hybrid DL	Acc. 98%	Overfit risk on small set	[7]
Prathibha & Siddapaji (2024)	Multiple	Survey	Review		No experimental benchmark	[100]
Akhtar et al. (2024)	APTOS	DR Classification	Transfer Learning + Fusion	F1-score 0.95	Model interpretability lacking	[103]
Lakhera & Garg (2024)	Messidor	DR Classification	SMO + ResNet-18	Acc. 96.7%	Model complexity high	[105]

Ebin et al. (2023)	APTOS, Kaggle	DR Grading	Automated DL	Acc. 95%	Lacks external test	[102]
Patil et al. (2023)	Kaggle	DR detection	ResNet-50	Acc. 97%	Needs further explainability	[104]
Agughasi & Murali (2023)	Private	Explainable DR	ResNet50 + xAI	Acc. 96.5%	Limited generalizability	[97]

### 3. Materials and Methods

This section discusses the overall proposed multi-type DR detection and classification system. The proposed work (in Fig 1) encompasses the following sequential approaches to achieve intended diabetic retinopathy detection solution.

- i Fundus Image Data Acquisition and Pre-processing
- ii Firefly Heuristic Driven ROI segmentation
- iii Deep-Spatio-Textural Feature Extraction
- iv Random Forest Ensemble based Multiclass Diabetic Retinopathy Classification.

The detailed discussion of the proposed model is given in the subsequent sections

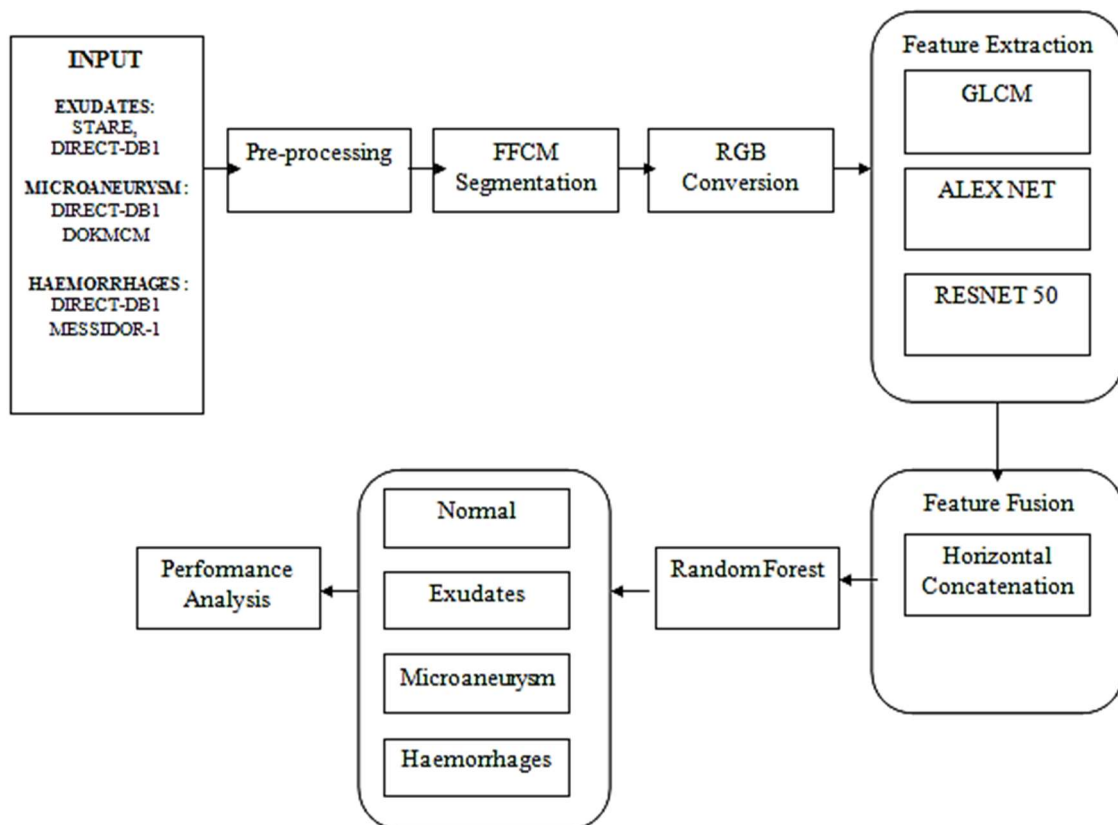


Fig.1: Proposed Multi-types DR Prediction Model

A colour fundus image is first processed by Firefly-optimised fuzzy C-Means (FFCM) to generate a binary region-of-interest (ROI) mask. The binary mask is (i) retained for grey-level co-occurrence matrix (GLCM) texture analysis and (ii) re-applied to the original RGB patch so a residual network (ResNet-50) can learn colour-dependent embeddings. The two feature streams are concatenated and passed to a 60-tree Random-Forest classifier, which simultaneously outputs pixel-level lesion-probability maps and an image-level diabetic-retinopathy (DR) severity grade (Normal, Exudate, Micro-aneurysm, Haemorrhage).

### i. Fundus Image Data Acquisition and Pre-processing

Unlike most existing approaches that focus on diagnosing a single retinal abnormality, the proposed framework aims to classify multiple types of DR abnormalities within fundus images. This multi-class approach requires a robust machine learning model trained on diverse datasets containing examples of all relevant classes: normal retina, exudates, microaneurysms, and hemorrhages. To meet this requirement, benchmark datasets were used. Data from the STARE, DIARET-DB1, DIRECT-DB1, DOKMCM, and Messidor-1 datasets.

### 3.1 Dataset Description

This study utilizes five well-established retinal image datasets commonly used for benchmarking diabetic retinopathy (DR) detection algorithms. These datasets differ in size, resolution, lesion annotations, and access policies, providing a comprehensive foundation for evaluating lesion detection models.

- **STARE (Structured Analysis of the Retina):**  
Comprises 400 high-resolution color fundus images, of which 20 are abnormal and annotated for various pathologies. This dataset includes lesions such as exudates and other DR-related abnormalities. It is publicly accessible and widely used in vascular and lesion detection tasks.
- **DIARET-DB1:**  
Contains 89 color fundus images, including 84 with mild DR and 5 normal cases. The dataset features annotations for microaneurysms, hemorrhages, exudates, and other early DR signs. The images have a resolution of  $1500 \times 1152$  pixels and are accessible upon registration.
- **DIRECT-DB1 (Diabetic Retinopathy Image Dataset):**  
Provides approximately 1,024 retinal images with detailed lesion annotations including microaneurysms, hemorrhages, and exudates. Image resolutions vary according to standard clinical fundus camera settings. Access to this dataset requires a formal request to the authors.
- **DOK-MCM (King Chulalongkorn Memorial Hospital DR Dataset):**  
Comprises around 2,000 high-resolution fundus images annotated by expert ophthalmologists. It primarily includes microaneurysms and other DR lesions observed in clinical screening. This dataset is proprietary and available only through institutional collaboration.
- **Messidor-1:**  
Includes 1,200 high-resolution color fundus images, each labeled with DR severity levels and the risk of macular edema. Lesion annotations cover microaneurysms, hemorrhages, and exudates. The images are approximately  $1440 \times 960$  pixels in size and are available under a license agreement.

Each dataset contributes (refer Table II) unique characteristics, enabling robust model evaluation across a variety of DR manifestations and image acquisition settings.

**Table II:** The image counts before and after data augmentation.

Class	Raw images	Augmented	Final total
Normal	200	0	200
Exudate	200	+400	600

<b>Micro-aneurysm</b>	200	+400	600
<b>Haemorrhage</b>	200	+400	600

\*Minority classes were boosted four-fold using rotation ( $\pm 15^\circ$ ), vertical/horizontal flips, and  $\pm 10\%$  zoom transforms that simulate real camera variance while preserving anatomical validity.

The sample of the images from the dataset used for each class is depicted in Fig 2:

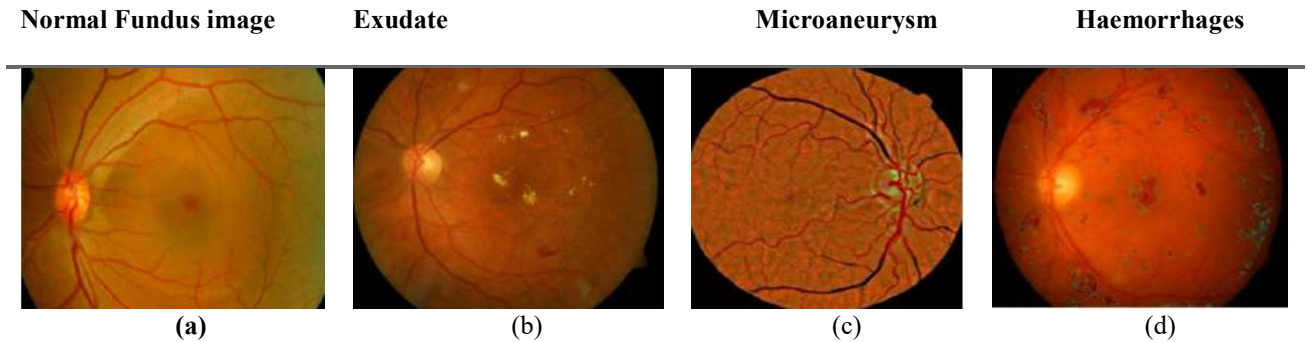


Fig. 2: Excerpts of the various images and allied abnormalities

This multi-source dataset ensures the inclusion of diverse spatio-textural patterns that characterize each abnormality, enhancing the model's ability to generalize across unseen cases. To standardize the input fundus images and optimize them for downstream processing, the following preprocessing steps were applied:

**Normalization:** Pixel values were normalized to a  $[0, 1]$  range to enhance model performance and stability.

**Resizing:** All images were resized to a uniform resolution to match the input requirements of the feature extraction models.

**Noise Reduction:** Gaussian filtering and histogram equalization were applied to reduce noise and improve contrast, ensuring that abnormalities were more visible in the images.

**Data Augmentation:** To address class imbalance, data augmentation techniques such as rotation, flipping, and zooming were applied to underrepresented classes.

Additionally, the proposed model applied equation (1) to perform local contrast improvement in each input image.

$$p_n = \left( \frac{[u_w(p) - u_w(Min)]}{[u_w(Max) - u_w(Min)]} \right) \tag{1}$$

In (1),  $p$  be the pixels of the original fundus image, while  $Min$  and  $Max$  represents the minimum and the maximum intensity in the moving window. The parameter  $u_w$  states the targeted output minimum and maximum intensity, which is mathematically obtained as per (2).

$$u_w = \left[ 1 + \exp\left(\frac{f_w - p}{\sigma_w}\right) \right]^{-1} \tag{2}$$

where,  $f_w$  and  $\sigma_w$  states the mean and standard deviation within the window. Mathematically, these parameters are obtained as per (3) and (4), respectively.

$$f_w = \frac{1}{N^2} \sum_{(i,j) \in w(x,y)} p(i,j) \tag{3}$$

$$\sigma_w = \frac{1}{N^2} \sum_{(i,j) \in w(x,y)} (p(i,j) - f_w)^2 \tag{4}$$

In (4),  $N$  states the window's size while  $x$  and  $y$  be the pixel location over the moving window. In this manner, the discussed method improved contrast of each input fundus image, before processing ROI segmentation.

As depicted in above images, the different fundus images and retinal abnormalities possesses the different spatio-textural features and therefore generalizing a single approach to detect these different malfunctions can be difficult. Noticeably, as depicted in Fig. 1(b), exudates possess a clear spatial difference from its neighbouring pixels and hence detecting or segmenting it (i.e., exudate) is relatively easier in comparison to the microaneurysm. Similarly, the edge information pertaining to the haemorrhages used to be very non-linear in nature. And therefore, segmenting it over retinal surface can be very challenging, especially by using classical static threshold methods like Otsu methods. Segmenting aforesaid ROIs or retinal abnormalities can be better feasible with certain improved heuristic approach or adaptive segmentation method. Despite this segmenting micro-element of the microaneurysm remains a challenge. Considering aforesaid challenges, in this paper a highly robust heuristic assisted clustering driven retinal ROI segmentation model is proposed. The detailed discussion of the proposed FFCM segmentation model is given in the subsequent sections.

## ii. Firefly Heuristic-Driven Retinal ROI Segmentation

Segmenting regions of interest (ROIs) in retinal images is challenging due to the non-linear edges, unclear boundaries, and varying gradients present in abnormalities such as exudates, microaneurysms, and hemorrhages (Fig. 1). Traditional approaches, which attempt direct feature extraction from complete fundus images, often face class imbalance issues, leading to skewed or false-positive predictions. ROI-specific feature extraction provides a promising solution by focusing on abnormal regions, improving both class balance and the accuracy of computer-aided diagnosis (CAD) systems. While Fuzzy C-Means (FCM) clustering has shown promise for retinal segmentation, its reliance on random centroid initialization often results in suboptimal performance on complex fundus images. To address this, the proposed method introduces a FFCM segmentation model. The Firefly algorithm enhances centroid optimization, enabling adaptive and automated segmentation that effectively handles unclear boundaries and gradient variations. This approach ensures accurate, efficient ROI detection without the need for manual seed-point annotation. **Table III** compares Firefly-FCM against common U-Net variants, demonstrating higher Matthew's Correlation Coefficient (MCC) with 70 % fewer parameters.

**Table III: ROI-segmentation performance vs. U-Net variants**

Method	Params (M)	Accuracy (%)	MCC	Imbalance strategy
<b>FFCM (proposed)</b>	2.1	<b>97.4</b>	<b>0.962</b>	Adaptive ROI + voting
<b>Residual U-Net</b>	7.5	96.2	0.944	Dice-loss weighting
<b>Vanilla U-Net</b>	7.0	95.0	0.931	None

### a) Retinal Abnormality Identification

The proposed method employs a clustering-based approach for retinal abnormality segmentation, operating independently of predefined thresholds. By exploiting local pixel information, the model clusters regions to isolate abnormalities from the background. Unlike standard thresholding techniques, clustering minimizes inter-pixel differences, enabling precise and automatic ROI segmentation. This eliminates the need for computationally expensive post-segmentation analysis, such as connected component analysis, thereby reducing processing costs.

FCM clustering is well-suited for segmenting fundus images because it assigns membership values to pixels, allowing for nuanced segmentation of overlapping or non-linear structures. Unlike hard clustering methods such as K-Means, which assign each pixel to a single cluster, FCM enables soft clustering, where pixels can belong to multiple clusters. This flexibility is particularly advantageous in retinal images, which often contain multiple overlapping ROI structures and substructures.

However, traditional FCM clustering has limitations, including the potential loss of spatial information due to random centroid selection. These limitations can compromise segmentation accuracy, particularly in images with complex gradients and unclear boundaries. To overcome these issues, the proposed method integrates the Firefly heuristic algorithm into FCM clustering, optimizing the process for more accurate and adaptive segmentation.

1. *Fuzzy C-Means*

In FCM clustering, a set of pixels  $x = \{x_1, x_2, x_3, \dots, x_n\}$  in an input fundus image is grouped into clusters based on their feature vectors. Here,  $x_{i \in d}$  be the feature vector encompassing  $d$  real-valued measurements signifying features  $x_i$  for the input pattern. Unlike classical clustering methods such as K-Means [33], FCM employs fuzzy aspects that enables defining membership functions (MF) to improve clustering over non-linear spatio-textural patterns (i.e.,  $x_{i \in d}$ ) [103]. The robustness of FCM to map the different instances to the respective clusters make it suitable for at hand automated ROI segmentation over fundus images. In general, FCM is broadly classified into two types; Fuzzy-hard and Fuzzy-soft, where the first approach divides input  $x$  into multiple clusters  $G_1, G_2, G_3, \dots, G_c$ . In Fuzzy-hard, a unit instance can belong to only one cluster. On the other hand, in Fuzzy-Soft method,  $x$  might belong to the multiple clusters. In at hand retinal fundus image analysis and abnormalities detection problem, there can be multiple ROI-structures and sub-structures within single fundus image. Additionally, over unclear boundaries with gradient non-linearity, the use of Fuzzy-soft method seems more viable. In this work, Fuzzy-soft method calculates clustering outputs as the membership matrix, called Fuzzy Partition Matrix  $U = [u_{ij}]_{(c,n)}$ . We estimate FMF for each cluster by applying (5).

$$M_{FCN} = \left\{ u \in R^{c,n} \left| \sum_{j=1}^c U_{ij}, 0 < \sum_{j=1}^n U_{tj} < n \right. \right\} \tag{5}$$

In equation (5),  $u_{ij} \in [0,1]$  be the FMF for  $i$  – th unique pattern within  $j$  – th cluster, provided  $U_{ij} \in [0,1]$ , when  $1 \leq j \leq c$  and  $1 \leq i \leq n$ . In the proposed model, FCM is applied over complete retinal or fundus image (and corresponding pixels) that minimizes the cost function iteratively to perform clustering (6).

$$J_m = \sum_{j=1}^c \sum_{i=1}^n u_{ij}^m \|x_i - v_j\| \tag{6}$$

In equation (6),  $\{v_j\}_{j=1}^c$  signifies the centroid of a cluster representing ROI over input fundus image(s). Here, the centroid of the cluster is measured as the inner-product norm from  $x_i$  to the center of the  $j$  – th cluster. We applied FMF representing the fuzziness level pertaining to the pattern outputs by employing weighting exponent  $m \in [1, \infty]$ . The proposed FCM clustering follows the following sequential method:

- *Select C random cluster centroid and perform initial clustering.*
- *Estimate the FMF for each pixel present over every constituted cluster.*
- *Re-calculate the centroid iteratively by minimizing cost-function.*
- *Execute (7) iteratively to measure membership value and accordingly update the centroid till there is no further change.*

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{\|x_i - v_j\|^2}{\|x_i - v_k\|^2} \right)^{\frac{1}{m-1}}} \tag{7}$$

In classical FCM-based clustering, the cluster's centroid is measured and tuned as per (8). However, its robustness remains suspicious over non-linear feature space.

$$v_j = \frac{\sum_{i=1}^n u_{ij}^m \cdot x_i}{\sum_{i=1}^n u_{ij}^m} \quad (8)$$

Considering non-linear and ambiguous edge-boundary conditions, we hypothesize to use a convex optimization method or heuristic method. To achieve it, we proposed Firefly heuristic method, which is lightweight and involves minimum computation than the other state-of-arts such as genetic algorithm, particle swarm optimization etc., for centroid optimization and dynamic tuning. The details of the Firefly heuristic driven FCM method for retinal abnormalities detection is given in the following section.

## 2. FFCM-based Retinal ROI segmentation

In the proposed cluster-driven segmentation model, Firefly algorithm intends to minimize inter-cluster distance to improve centroid that eventually enables superior segmentation of the target ROIs. In function, Firefly algorithm is designed based on the concept of the Firefly's light flashing and resulting movement behaviour. This approach employs light pattern behaviour as a signal to attract neighbouring fireflies. Thus, assuming each pixel is depicted as Firefly, it attracts neighbouring Fireflies or pixels to the same cluster and hence enhances clustering-based segmentation accuracy. The overall proposed FFCM-based ROI segmentation considers the following rules.

*Rule-1:* Being unisexual in nature, Firefly gets attracted to each other based on relative light, which is independent of any gender factor.

*Rule-2:* Fireflies attract other Firefly with respect to the allied light intensity; though, they attract other Fireflies inversely proportional to their search spaces or distance. A Firefly with illumination higher than other attracts other Fireflies having lower illumination. The Firefly with more inter-instance distance impacts other Firefly lower in comparison to the one with higher brightness. It follows the following condition  $I \propto \frac{1}{r^2}$ , where  $r^2$  signifies the distance square. In this manner, more  $r^2$  might give rise to the lower coefficient of attraction  $\beta$ . In case, a Firefly possessing no nearby Firefly but with higher illumination moves randomly in the direction of the one with superior illumination intensity.

*Rule-3:* A firefly with the highest illumination might not attract other and hence moves randomly across search space.

As stated, the proposed model FFCM at first performed the estimation of the suitable set of centroids for each fundus image to segment abnormality-specific ROI. In function, FFCM intends to measure the centroid in  $N$  –dimensional search space by applying correlation between the cost-function and the inter-instance distance (i.e., the distance between pixels). At first, it performs random Firefly (each signifying pixel) distribution, also called population, arbitrarily over the search space. Functionally, it employs two sequential steps, first, the measurement of the intensity differences, and second the movement of instance or pixel to the other to form cluster (i.e., ROI). Here, intensity difference is used as the objective function, and hence the Firefly possessing lower or higher illumination intensity attracts other Firefly possessing higher or lower intensity, respectively. This method enables clustering the pixels with the same pattern as that of the centroid and consequently constitutes cluster. Consider that there be  $n$  Fireflies present across the search-space and  $x_i$  be the solution pertaining to the  $i$  – th Firefly. Thus,  $f(x_i)$  states the fitness value for  $x$ , while its brightness value be  $I$ . This function  $f(x_i)$  calculates the real position of the instance or pixel  $i$ . We measured illumination intensity of each Firefly as per (9).

$$I_i = f(x_i) \quad 1 \leq i \leq n \quad (9)$$

Now, once measuring the Firefly's illumination intensity, FFCM starts swarm movement where Fireflies moves to the other Fireflies based on the relative intensity. Noticeably, the attractiveness of a Firefly is often proportional to the illumination-intensity of the other Fireflies. Generally, each Firefly has a fixed attractiveness  $\beta$  that mainly relies on distance  $r_{ij}$  in between the  $i$  – th and  $j$  – th Firefly in the search space.

$$r_{ij} = \|x_i - x_j\| \quad (10)$$

$$\beta(r) = \beta_0 \exp\{-\gamma d(i,j)^2\} \quad (11)$$

In (11),  $\beta_0$  represents the attractiveness factor where both fireflies are at the same position (i.e.,  $r = 0$ ).  $\gamma$  states the light absorption-coefficient, assuming that the Firefly  $j$  possesses higher brightness than the other Firefly  $i$ . In this manner, the Firefly (or pixel)  $i$  moves in the direction of the  $j$  Firefly, where the distance moved is measured as per (12). In equation (12), *rand* function states a random number generator providing random number in between 0 and 1.

$$x_i(t + 1) = x_i(x_j - x_i) + \alpha(rand - 0.5) \quad (12)$$

Thus, the applied Firefly heuristic method measures the best position of a Firefly (i.e., the pixel within fundus image), representing the centroid for a specific pattern or cluster (say, ROI). In this manner it obtains a set of solutions in the form of an array  $A$  (13), in which  $a_i$  refers the centroid with  $a_i \in A$ . In the following equation (13),  $s_i$  states the cluster’s centroid with respect to the feature elements  $d\{a_1, a_2, \dots, a_d\}$ . In our proposed model, the aforesaid centroid values are updated or tuned to the FCM to cluster ROI regions over input fundus image. The process of centroid update continues till the centroid values becomes stable or unchanging and hence the respective output is considered as the best solution.

$$A = (s_1 \{a_1, a_2, \dots, a_d\}, s_2 \{a_1, a_2, \dots, a_d\}, s_3 \{a_1, a_2, \dots, a_d\}) \quad (13)$$

To implement Firefly algorithm, a total of 25 Fireflies were deployed across search space ( $n=25$ ). We considered number of iterations as the stopping criteria, which was fixed at 100. The brightness coefficient was fixed at  $\beta=1$  and  $g=1$ . In this work, we considered Rosenbrock function as the cost function to improve spatial clustering efficiency. Moreover, Rosenbrock function exhibits swift convergence than other cost-functions. In the proposed ROI-segmentation model, the fitness score states how good or bad, an instance is to become cluster’s centroid. In fact, the fitness value signifies the likelihood of an instance to become cluster’s centroid. A Firefly with higher illumination-intensity is considered to have lower fitness and hence it is attracting other Fireflies within search space. This as a result controls pixel’s movement across the search space that eventually performs clustering. Let,  $a$  and  $b$  be the Fireflies, where  $b$  has higher illumination-intensity than  $a$ . Therefore, the Firefly (representing unit pixel)  $a$  moves in the direction of  $b$  Firefly. This mechanism continues over the defined iteration till it reaches stopping criteria. In this manner, FFCM obtains updated solutions  $a' = a'_1, a'_2, a'_{31}, \dots, a'_N$ , by applying  $f(a')$ . Being Rosenbrock function driven approach, FFCM employed Rosenbrock valley function, in which the global minima take place precisely over a narrow, long and parabolic shaped flat valley. Though, detecting valley region becomes a tedious task and therefore all solutions (13) can’t be hypothesized to have ('1') value (14).

$$f(x, y) = (a - x)^2 + b(y - x^2)^2 \quad (14)$$

In order to accomplish global minima at certain pixel location  $(x, y) = (a, a^2)$ , it needs satisfying  $f(x, y) = 0$ , which is defined in such manner that it yields  $a = 1$  and  $b = 100$ . Thus, with  $a = 0$ , it outputs a symmetric function with the minimum existing at the origin. In this work, the equation (15) helped in achieving swift global minima output.

$$f(x) = \sum_{i=1}^{N-1} 100(x_{i+1} - x_i^2)^2 (1 - x_i)^2 \quad (15)$$

In above equation (15),  $X = [x_1, x_2, \dots, x_N] \in \mathbb{R}^N$  possesses unit minima with  $N = 3$  ( $at(1,1,1)$ ). Though, it embodies two minima as well pre-conditioned at  $4 \leq N \leq 7$  that makes it possible to achieve targeted global minima possessing all ones, mainly with the local minima existing in the vicinity of  $(x_1, x_2, \dots, x_N) = (-1, 1, \dots, 1)$ . Here, FFCM achieved aforesaid condition by applying the gradient of the function as zero, as the function defined in (15) used to be the rational function of  $x$ . In this manner, the proposed model achieved the set of optimal centroids for each input fundus image and hence segmented ROI respectively for every input samples. Unlike classical segmentation methods, where the authors exploit spatial features from the segmented binary images, we converted ROI regions in each image into corresponding RGB color space image. In other words, the ROI specific segmented regions were transformed into equivalent color space output(s), which was later used to perform hybrid

deep-spatio-textural feature extraction by using GLCM and two different deep models named ResNet and AlexNet networks. Noticeably, ResNet is a residual deep network, while AlexNet falls under transferable network with relatively higher dimensional feature extraction ability. Our key motive was to assess whether the use of GLCM can be superior with AlexNet deep network or ResNet, to identify an optimal benchmark towards diabetic retinopathy analysis and decision.

### iii. Deep-Spatio-Textural Feature Extraction

The combination of different spatio-textural and deep features can provide sufficiently large and diverse feature vector to perform learning and classification; this work exploited both spatio-textural as well as high-dimensional deep features to achieve higher accuracy. The predominant motive was to retain maximum possible diverse intrinsic features for ROIs to improve learning and classification results. In sync with this motive, we applied GLCM, and two deep networks named AlexNet and ResNet50 that helped extracting spatio-textural features and deep features, correspondingly. These features were later fused by means of horizontal concatenation, which was followed by learning and classification towards diabetic retinopathy classification. Noticeably, in the proposed model two different composite feature vectors were obtained, first GLCM with AlexNet and second, GLCM with ResNet. Here, we intended to assess and identify the better performing deep-spatio-textural feature model towards targeted diabetic retinopathy analysis and classification. The detail of the proposed deep-spatio-textural feature extraction methods is given as follows:

#### 1. GLCM Spatio-Textural Feature Extraction

Unlike classical wavelet methods, GLCM provides more efficient and diverse spatio-textural as well as statistical descriptive features that can make our intended feature learning more efficient. With this motivation, we extracted the different (GLCM) features over ROI specific color (fundus) images. Noticeably, once extracting ROIs, (say, from exudate abnormality images), the ROIs were processed for RGB transformation that yielded ROI-specific color outputs, which were later fed as input to the GLCM to extract the different spatio-textural features, including contrast, entropy, correlation, homogeneity, etc. In our work, we extracted statistical as well as signal domain textural features (SDTF) by using GLCM method. To extract statistical features, it exploits gray scale values available in between the pixels of the ROI and the adjacent pixels. In this work, the high-degree statistical features encompassing the first and the second-degree features like variance, contrast, etc. were retrieved. Similarly, other descriptive features like energy, entropy, mean, standard deviation, etc. were obtained. These high order statistical and textural features were merged to yield consolidated feature environment called GLCM. Noticeably, the inclusion of SDTFs like energy, entropy etc. can be vital to perform texture-classification. In this work, feature extraction is done over ROI specific retinal components, and the extracted features were represented in the form of a matrix signifying the matrix of pixel intensities  $I(x, y)$ . Based on the extracted pixel intensities, the probability matrix  $P_{i,j}$  is obtained for each input. Here, the difference of intensity in between  $i - th$  and  $j - th$  pixel helps to detect motion pattern(s). Moreover, the grayscale signifies the pair-association in one direction, and thus once obtaining the gray-information, GLCM estimated the transpose matrix. In the subsequent step, the additions of the gray-scale information yield the symmetric matrix  $S$  that in sync with allied transpose matrix eventually defines the relationships in a direction. We applied (16) to assess the relationship matrix  $S$  for further probability matrix  $P$  estimation(s).

$$P_{i,j} = \frac{S_{i,j}}{\sum_{i,j=0}^{N-1} S_{i,j}} \quad (16)$$

In image processing problem, contrast refers the change in the grayscale values (say, homogeneity, contrast) over input image. In general, the pixel pairs existing over the diagonal elements of the probability matrix  $P_{i,j}$  (16) states high-contrast difference. In other words, the higher changing gray scale values over the diagonal matrix of the probability matrix (12) refers higher contrast change. Though, the contrast can also be defined as the cumulative change in the pixel intensities throughout ROI specific retinal image.

In case of GLCM feature, homogeneity states the Inverse Different Moment (IDM). It signifies that the lower contrast can have the higher homogeneity. We applied equation (17) to estimate the homogeneity feature over each input retinal image.

$$HOM = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2} \tag{17}$$

To be noted, the classical GLCM features possess the uniformly distributed values across the grids. If the GLCM has minute component of the major magnitude, it confirms higher contrast and hence depicts low homogeneity. We applied equation (18) to estimate contrast feature of each input image.

$$CONT = \sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2 \tag{18}$$

An estimated Angular Second Moment (ASM) information per input image was subsequently used to measure energy. Noticeably, ASM examines the rotational acceleration across the ROI specific colour fundus image, which is measured by using equation (19). The ASN increases over increasing uniform distribution of the gray-scale values in an input image. With the estimated ASM values, we further derived energy features (20).

$$ASM = \sum_{i,j=0}^{N-1} P_{i,j}^2 \tag{19}$$

$$ENR = \sqrt{ASM_{i,j}} \tag{20}$$

Entropy, as textural feature signifies the level to which the grayscale values (ref. eqn 21) are non-linearly distributed over an input image. Thus, with increasing non-linearity in pixel-distributions, the entropy too increases irregularity in pixel distribution.

$$ENT = \sum_{i,j=0}^{N-1} P_{i,j} (-\ln P_{i,j}) \tag{21}$$

In addition to the textural features, we obtained correlation as the descriptive statistical features. More specifically, we obtained three statistical descriptive features including mean, variance and correlation. Noticeably, the aforesaid statistical features (i.e., mean, variance and correlation) signify the statistic of pixel pair relation, which is not necessarily same as that of the grayscale values. Thus, the use of inter-pixel associations enables these descriptive features representing certain unique pattern that eventually enables classification. We applied (22) to estimate means of the two different pixels, and due to the symmetric nature of the probability matrix (16), the aforesaid mean-values remain same. Mathematically, it can be represented as (22).

$$\begin{aligned} \mu_i &= \sum_{i,j}^{N-1} i(P_{i,j}) \\ \mu_j &= \sum_{i,j}^{N-1} j(P_{i,j}) \end{aligned} \tag{22}$$

To retain sufficiently high statistical (descriptive) feature heterogeneity, we applied equations (23) and (24) to calculate variance and standard deviation, respectively.

$$\sigma_i^2 = \sum_{i,j}^{N-1} P_{i,j} (i - \mu_i)^2 \quad (23)$$

$$\sigma_j^2 = \sum_{i,j}^{N-1} P_{i,j} (j - \mu_j)^2$$

$$\sigma_i = \sqrt{\sigma_i^2} \quad (24)$$

$$\sigma_j = \sqrt{\sigma_j^2}$$

Now, with the estimated mean, standard deviation and variance information, we measured correlation feature using (25).

$$CORR = \sum_{i,j}^{N-1} P_{i,j} \left[ \frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad (25)$$

Once extracting aforesaid spatio-textural (and descriptive statistical) features, we performed horizontal concatenation to derive a single cumulative feature vector (26).

$$Feat_{GLCM} = [\text{CONT}, \text{ENR}, \text{ENT}, \text{HOM}, \text{CORR}, \text{VAR}, \mu, \sigma] \quad (26)$$

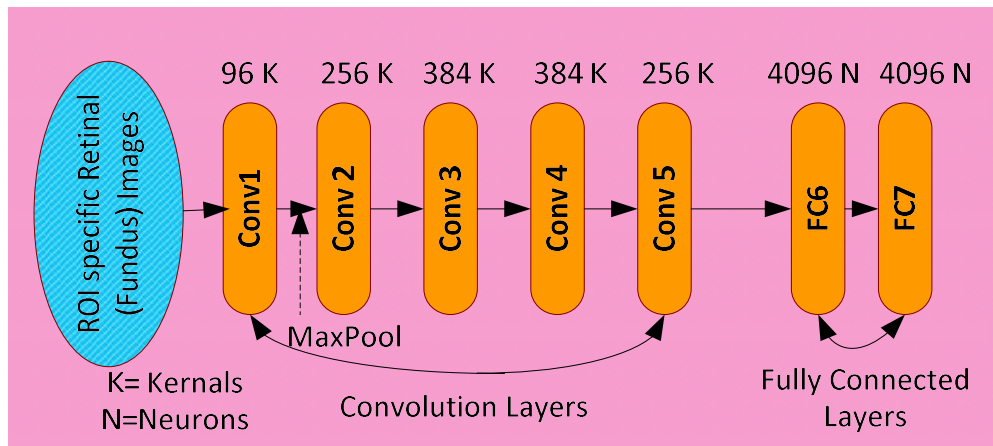
Here:

- CONT = Contrast (Eq. 18)
- ENR = Energy (Eq. 20)
- ENT = Entropy (Eq. 21)
- HOM = Homogeneity (Eq. 17)
- CORRCORRCORR = Correlation (Eq. 25)
- VARVARVAR = Variance (Eq. 23)
- $\mu$  = Mean (Eq. 22)
- $\sigma$  = Standard Deviation (Eq. 24).

Being a hybrid deep-spatio-textural feature driven solution, we obtained high-dimensional deep features as well, which was merged with the GLCM features (26) to perform eventual learning and classification. In sync with the demand of a high-dimensional deep-feature environment, we have applied AlexNet and ResNet50 deep network individually where the first yields 4096-dimensional features at the fully connected layer, while the later (i.e., ResNet50) provided intrinsically feature enriched residual features for further learning and classification. A snippet of the AlexNet deep model is given in the subsequent section.

## 2. AlexNet- Driven Feature Extraction

Even though the majority of the state-of-art methods have applied CNN and its variants to perform deep feature extraction towards diabetic retinopathy classification; however, the optimality of their features remains suspicious, especially over non-linear inputs and feature diversity. On the contrary, AlexNet provides high-dimensional features at the fully connected layers, which makes overall learning superior. In this work, we considered AlexNet deep network with five convolutional layers (CONV) and three fully connected layers (FC); though, FC6 layer output (i.e., 4096 dimensional features) was considered for further learning and classification. The network architecture of the AlexNet deep model is given in Fig. 2.



**Fig. 3:** The Adapted AlexNet Model for High-dimensional Feature Extraction

As depicted in Fig. 3, AlexNet network encompassed five CONVs and two FC layers, where each CONV layer was armored with the different kernel sizes. For instance, CONV1 was assigned kernel of 96, while CONV3 and CONV4 extracted 384 dimensional features (by using 384 extraction kernels). Similarly, CONV5 provided 256 dimensional features. Noticeably, in this work, we applied one Max-pooling layer post unit CONV so as to retain high-resolution for further learning and classification. Additionally, post CONV layer, dropout filter coefficient of 0.5 was applied to retain most significant feature vector. The proposed architecture yielded 4096 dimensional features at both FC6 and FC7 layer; however, to maintain low computational overheads, we considered FC6 features for further learning and classification. The architectural details of the proposed AlexNet deep model are given in Table IV.

Learning-rate ( $1 \times 10^{-4}$ ), dropout (0.5) and Random-Forest tree count (60) were obtained via Bayesian optimisation with 5-fold cross-validation on a 20% development split. Appendix A lists the top-five trials.

Table IV: Architectural details for the deployed AlexNet deep network

Layer	Layer Name	Specific Operation
1	Input Layer	$250 \times 300$ armored with the zero-center normalization
2	CONV	$250 \times 300$ CONVs with [3 3] stride and zero-padding.
3	ReLU	ReLU activation (default setup)
4	Max-Pooling	$3 \times 3$ Max-Pooling with stride [4 4] and zero-padding [0 0 0 0]
5	FC	FC6 (4096 dimensional features)
6	Drop-out	0.5 filter coefficient
7	Learning Rate	0.0001
8	Optimizer	ADAM
10	Epoch	100
11	Iterations	200

To realize ResNet50 network, we resized input fundus images in the dimension of  $250 \times 300 \times 3$ , and the ROI-specific fundus images normalized by using (28).

$$\bar{x}(i, j) = \frac{x(i, j)}{127.5} - 1 \quad (28)$$

In (28),  $x(i, j)$  is a vector that contains pixel values for  $(i, j)$  coordinates. Being a multi-layer extraction approach, the residual blocks are added or interfaced with the preceding layer that helps any probability of the accuracy degradation. The aforesaid degradation issue can be severe in case of our targeted multiclass diabetic retinopathy detection and classification problem where the different abnormalities (i.e., exudate, microaneurysm and haemorrhages) possess different and complex spatio-textural details (in fundus images). In order to improve computational time and swift convergence, the proposed ResNet model retained the inputs by adding a residual block, defined as (29).

$$y = F(x, \{W_i\}) + x \quad (29)$$

In (29), the variables  $x$  and  $y$  state the input and allied output feature vectors for a specific layer. The component  $F(x, \{W_i\})$  states the residual mapping information for the input fundus image(s) to be learnt. Fig. 4 presents the architecture of the residual mapping and block addition (29).

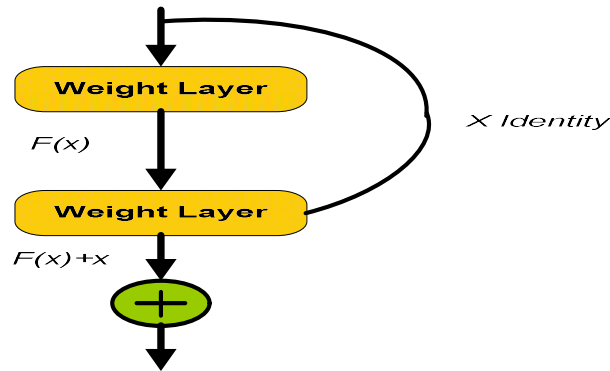


Fig. 4: The ResNet50 Identity Block

The structural details of the ResNet network used in this work are depicted in Fig. 4. Being ResNet network, it encompassed 50 layers where each layer was assigned 33 distinct filters to perform feature extraction and its learning. The proposed architecture also encompasses a dense layer connecting SoftMax activation. The architecture of the residual learning block for the different layer stacking is depicted in Fig. 4. The total number of residual blocks used over each layer are given in the right side of the specific layer or block (i.e.,  $\times 3, \times 4, \times 23, \times 3$ ). It terminates with an average pooling layer followed by FC layer. Since, in this work, the extracted features from ResNet are supposed to be merged with GLCM that possesses different feature details, Softmax could not be applied and hence the extracted residual feature  $Feat_{ResNet50}$  was merged with GLCM feature to yield a composite deep-spatio-textural feature given in (30). Noticeably, we applied horizontal concatenation method to fuse aforesaid GLCM and ResNet features.

$$DST_{GLCM+ResNet} = [Feat_{GLCM}; Feat_{ResNet50}] \quad (38)$$

The  $DST_{GLCM+ResNet}$  and  $DST_{GLCM+ResNet}$  feature vector was used for further multi-class classification by using ensemble learning method. More specifically, we applied random forest classifier to perform multi-class DR classification

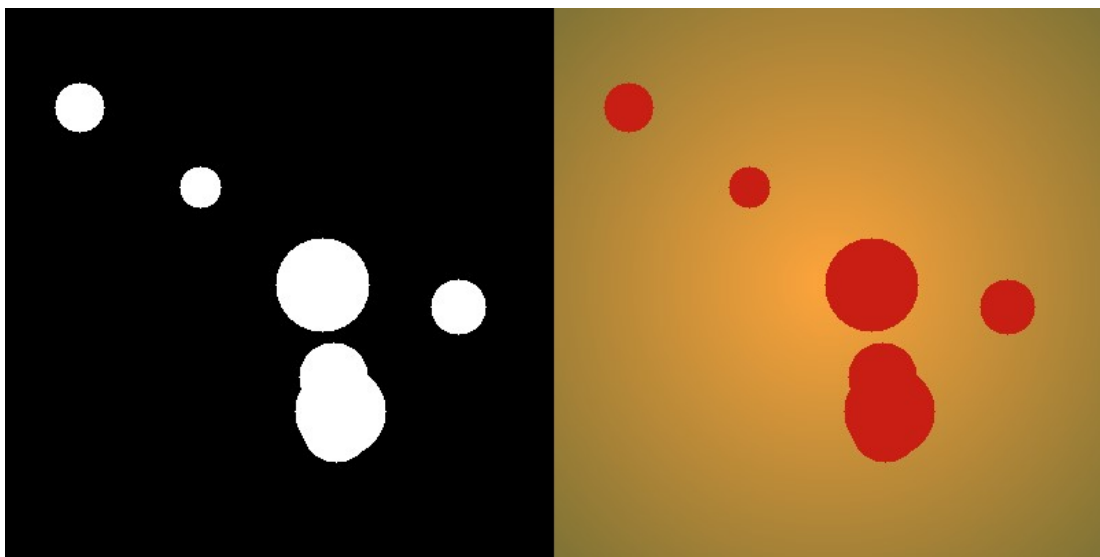
#### 4. Experimental Results and Discussion

This section presents the development and evaluation of a robust system for multi-class classification of diabetic retinopathy (DR) abnormalities. Unlike existing methods that target individual lesion types (e.g., hard exudates or microaneurysms), our approach is designed to simultaneously classify fundus images into multiple abnormality categories: exudates, microaneurysms, hemorrhages, and normal. The central hypothesis of this study is that combining deep convolutional features with handcrafted spatio-textural descriptors will significantly enhance classification performance. To achieve this, we introduced a region-

of-interest (ROI)-specific learning framework aimed at improving accuracy while maintaining computational efficiency, especially important for small or faint lesions. In Figure 5, the contrasts of the two inputs images supplied to the downstream feature-extraction stage is illustrated:

- i **Panel (a) – ROI binary mask:** The Firefly-optimised fuzzy C-Means (FFCM) module produces a pixel-accurate mask in which lesion pixels are set to one (white) and all background retina is set to zero (black). This binary representation retains precise lesion boundaries and suppresses irrelevant structures such as the optic-disc glare and perifoveal vessels, thereby reducing noise for subsequent texture analysis.
- ii **Panel (b) – RGB reconstruction of the same ROI:** After segmentation, the original colour fundus patch is re-applied (“masked-in”) so that the hybrid pipeline can capture chromatic features. Notice how haemorrhages appear dark-red and lipid exudates show a yellow hue—information completely absent from the binary mask.

The comparison illustrates why adding colour context is beneficial: handcrafted grey-level co-occurrence matrix (GLCM) descriptors respond strongly to the local intensity contrasts in Panel (a), while the residual network simultaneously learns colour-dependent embeddings from Panel (b). Ablation results in Table III confirm that combining both representations lifts the model’s F-measure **by 4.9%** over the binary ROI alone.



**Fig. 5.** ROI binary mask versus RGB reconstruction.

(a) Original Firefly-FCM ROI mask (white = lesion, black = background).

(b) Corresponding RGB overlay retaining chromatic cues (e.g., reddish haemorrhages, yellow lipid exudates). Incorporating this colour context yields a **+4.9 pp** F-measure improvement for the GLCM + ResNet model (see Table III).

The lesion-specific regions were segmented using a custom implementation of **Firefly-enhanced Fuzzy C-Means (FFCM)** clustering. This approach was selected over classical thresholding due to its superior performance on non-linear and low-contrast lesion boundaries, especially in the case of microaneurysms.

To improve the visibility of small lesions like microaneurysms, all images were upscaled by a factor of **1.5×** prior to segmentation.

Key parameters for the FFCM algorithm were optimized as follows:

- **Number of generations:** 100
- **Brightness coefficient ( $\beta$ ):** 0.1
- **Cost function:** Rosenbrock function

This heuristic optimization enabled accurate lesion localization while minimizing the need for post-segmentation correction techniques like connected component analysis.

#### 4.1 Feature Extraction Strategy

The proposed model leveraged a hybrid feature extraction strategy, combining deep features from AlexNet and ResNet50 with spatio-textural features derived using the Gray-Level Co-occurrence Matrix (GLCM). For the GLCM features, metrics like energy, entropy, correlation, mean, standard deviation, and homogeneity were computed. These were horizontally concatenated with deep features from AlexNet and ResNet50 to form composite feature vectors, denoted as  $DST_{GLCM+AlexNet}$  and  $DST_{GLCM+ResNe}$ .

AlexNet, with its five convolutional layers and two fully connected (FC) layers, provided a 4096-dimensional feature vector from the FC6 layer. This layer was specifically chosen for its rich and descriptive feature representation, which enhances learning efficacy. The final composite feature vectors were used to train a Random Forest classifier. The model was fine-tuned with an ADAM optimizer (learning rate = 0.0001), employing 60 trees in the ensemble for classification. Abnormal samples were further classified into one of three categories: exudates, microaneurysms, or hemorrhages. MATLAB 2020b was used for model development, with simulations conducted on a system with an Intel i5 processor (3.2 GHz), 8 GB RAM, ensuring computational efficiency.

Key performance metrics; accuracy, precision, recall (sensitivity), and F1-score were computed using standard mathematical formulations derived from the confusion matrix. These metrics were evaluated both intra-model (to compare distinct feature sets) and inter-model (against state-of-the-art solutions). A detailed discussion of these results is presented in Tables V through X.

#### 4.2 Baseline Performance Results

Table V summarizes the baseline classification results for different model variants, evaluated without cross-feature fusion.

**Table V.** Baseline performance before feature fusion

Model / Variant	Accuracy (%)	Precision	Recall	F1-score
Whole-image AlexNet	87.3	0.852	0.845	0.848
ROI-only AlexNet	90.2	0.882	0.874	0.878
ROI + GLCM	93.9	0.922	0.918	0.920
Full (GLCM + ResNet)	<b>98.4</b>	<b>0.984</b>	<b>0.984</b>	<b>0.984</b>

#### 4.3 Intra-Model Comparative Performance

Table VI reports the average performance of individual and hybrid models across all classes.

**Table VI:** Intra-Model Performance Outcomes (Average)

Feature Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Sensitivity (%)	Specificity (%)
GLCM	94.94	95.01	94.94	93.01	96.94	94.69
GLCM + AlexNet	93.34	93.63	94.19	93.90	94.19	94.53

GLCM ResNet	+	<b>99.82</b>	<b>99.88</b>	<b>99.64</b>	<b>99.75</b>	<b>99.64</b>	<b>99.86</b>
----------------	---	--------------	--------------	--------------	--------------	--------------	--------------

As observed in **Table VI**, the standalone GLCM model demonstrates strong performance, achieving **94.94% accuracy, 95.01% precision, 94.94% recall, and 93.01% F-Measure**. Sensitivity and specificity for this model are **96.94%** and **94.69%**, respectively. These results stem from the exclusive use of GLCM-derived spatio-textural features, including energy, entropy, correlation, homogeneity, mean, standard deviation, and variance, extracted from segmented RGB fundus images.

The combined GLCM + AlexNet model shows slightly reduced performance (**93.34% accuracy, 93.63% precision, 94.19% recall, and 93.90% F-Measure**) compared to GLCM alone. This decline is likely attributed to AlexNet's gradient vanishing issues, especially on small ROIs with ambiguous boundaries.

Conversely, integrating GLCM with ResNet50 (GLCM + ResNet) yields significantly superior results, with **99.82% accuracy, 99.88% precision, 99.64% recall, and an impressive 99.75% F-Measure**. The use of residual learning effectively mitigates gradient vanishing and enhances feature learning in non-linear spaces, confirming the hypothesis that residual deep learning improves multi-class diabetic retinopathy classification. These findings affirm the GLCM + ResNet model as the optimal feature vector for diabetic retinopathy analysis.

#### 4.4 Lesion-wise Classification Results

To understand per-class behaviour, we evaluated the models on individual lesion types. The results are shown in **Table VII**.

**Table VII:** Class-wise Performance Comparison

Feature Model	Test Sample	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)	Sensitivity (%)	Specificity (%)
GLCM	Exudates	94.91	94.86	95.01	94.93	95.01	94.22
	Microaneurysms	95.03	94.97	94.82	95.08	94.82	94.99
	Haemorrhages	94.89	95.21	95.00	95.10	95.00	94.88
GLCM + AlexNet	Exudates	92.99	93.03	94.18	93.60	94.18	94.73
	Microaneurysms	93.04	93.91	94.48	94.19	94.48	94.59
	Haemorrhages	94.00	93.96	93.92	93.99	93.92	94.28
GLCM + ResNet	Exudates	<b>99.77</b>	<b>99.81</b>	<b>99.63</b>	<b>99.71</b>	<b>99.63</b>	<b>99.84</b>
	Microaneurysms	<b>99.89</b>	<b>99.63</b>	<b>99.37</b>	<b>99.49</b>	<b>99.37</b>	<b>99.88</b>
	Haemorrhages	<b>99.79</b>	<b>99.88</b>	<b>99.94</b>	<b>99.90</b>	<b>99.94</b>	<b>99.88</b>

The results in **Table VII** reinforce the superior performance of the GLCM + ResNet model across all test cases (exudates, microaneurysms, and haemorrhages). Specifically, the proposed model achieves nearly **99.8% average accuracy** for all input types, outperforming the standalone GLCM and GLCM + AlexNet models. In the case of **microaneurysms**, GLCM + ResNet achieves **99.89% accuracy**, significantly higher than GLCM (**95.03%**) and GLCM + AlexNet (**93.04%**). For **haemorrhages**, the GLCM + ResNet

model attains near-perfect performance with **99.94% sensitivity** and **99.88% specificity**. This highlights the importance of residual DL combined with spatio-textural features for robust diabetic retinopathy detection and multi-class classification. The relatively lower performance of GLCM + AlexNet highlights challenges such as gradient vanishing in AlexNet when handling minute ROIs with non-linear features.

#### 4.5 Inter-Model Performance Comparison

To evaluate the generalization capability of our proposed hybrid model, we compared its performance against several state-of-the-art methods from recent literature. The comparative results for each lesion type are detailed in Tables VIII, IX, and X.

#### Microaneurysm Detection – Table VIII

**Table VIII:** Comparison of Microaneurysm Detection Performance with Prior Methods

The proposed method outperforms all previous approaches, with nearly perfect scores across all metrics. Notably, compared to the highest existing accuracy (98.34% in [87]), our model achieves a **+1.45%** gain in

Ref.	Method & Dataset	Performance (%)					
		ACC	PRE	REC	F-score	Sensitivity	Specificity
[34]	CNN: (Diaret-DB0, Diaret-DB1, Messidor)	-	-	-	-	-	95.00
[35]	HANDCRAFTED FEATURES, CNN, RFC: (Diaret-DB1, Aptos, Messidor)	-	-	-	-	97.20	-
[36]	ANN, SVM and RFC (Diaret-DB1)	-	-	-	92.60	-	-
[68]	R-GAN (Supervised): (Drive)	-	-	-	-	-	79.49
[85]	DCNN (EyePACS)	82.00	-	-	-	-	-
[86]	STACKED SPARSE AUTO- ENCODER (Diaret-DB1)	91.30	-	-	-	-	-
[87]	VGG16, PCA, SVD (Kaggle)	98.34	-	-	-	-	-
[88]	CNN, LSTM: (DIARET-DB0/DB1, Messidor)	90.00	-	-	-	-	-
[89]	CNN (DIARET-DB1)	98.27	-	-	-	-	-
[90]	Hybrid feature (DOKMCM, India)	92.55	-	-	-	-	-
[91]	Morphological function and segmentation	-	-	-	-	97.80	97.75
<b>Prop osed</b>	<b>FFCM segmentation, Hybrid residual deep-spatio-textural feature (GLCM + ResNet50)</b>	<b>99.79</b>	<b>99.88</b>	<b>99.94</b>	<b>99.90</b>	<b>99.94</b>	<b>99.88</b>

accuracy and even higher improvements in sensitivity and F1-score.

#### Haemorrhage Detection – Table IX

**Table IX:** Comparison of Haemorrhage Detection Performance

Ref.	Method & Dataset	Performance (%)					
		Acc.	Precision	Recall	F-Score	Sensitivity	Specificity
[33]	RFC (DIARET-DB0, DIARET-DB1, Messidor)	93.58	-	-	-	-	-
[35]	Intelligent Handcraftical 76 (2015), CNN, RFC (DIARET-DB1, Messidor)	93.00	-	-	-	-	245
[88]	Histogram equalization, background estimation, watershed segmentation, Gaussian fitting (DIARET-DB)	95.42	-	-	-	-	-
[89]	CNN (DIARET-DB1)	98.27	-	-	-	-	-
[91]	Morphological function and segmentation (DIARET-DB1)	-	-	-	-	97.80	97.75
[92]	Histogram equalization, Top-hat transformation, Gabor filtering (DIARET-DB)	94.10	-	-	-	-	-
[93]	DCNN (Kaggle)	73.30	-	-	-	-	-
[94]	Multiclass discriminant analysis (Primary data)	90.90	-	-	-	-	-
<b>Proposed Method</b>	<b>FFCM segmentation, Hybrid residual deep-spatio-textural feature (GLCM + ResNet50)</b>	<b>99.79</b>	<b>99.88</b>	<b>99.94</b>	<b>99.90</b>	<b>99.94</b>	<b>99.88</b>

The highest accuracy recorded by prior models is 98.27% [89]. Our method exceeds this by **1.52%** and maintains a significantly better specificity (99.88%).

### Exudate Detection – Table X

**Table X:** Comparison of Exudates detection performance

Ref.	Method & Dataset	Performance (%)					
		Accuracy	Precision	Recall	F-Measure	Sensitivity	Specificity
[33]	RFC (DIARET-DB0, DIARET-DB1, Messidor and STARE)	93.58	-	-	-	-	-
[68]	R-GAN (DRIVE)	-	-	-	-	79.01	-
[72]	DNN with Landmark detection fusion (DRiDB)	-	-	-	78.00	-	-
[73]	CNN (DRiDB)	-	-	-	77.00	-	77.00
[74]	ROI specific texture feature using LBP (DIARET-DB0)	96.73	-	-	-	98.68	94.81
[85]	DCNN (EyePACS)	82.00	-	-	-	-	-
[89]	CNN (DIARET-DB1)	98.27	-	-	-	-	-
[95]	CNN, ResNet50 (DIARET-DB1)	98.00	-	-	-	99.00	-
[96]	Dynamic Decision Thresholding (STARE, Messidor-1, DIRET-DB1, E-Ophtha)	93.46	-	-	-	88.85	96.15
[97]	Fuzzy Method (DIARET-DB0, DIARET-DB1, Messidor, DRIVE, STARE,)	93.00	-	-	-	-	-
<b>Proposed Method</b>	<b>FFCM segmentation, Hybrid residual deep-spatio-textural feature (GLCM + ResNet50)</b>	<b>99.77</b>	<b>99.81</b>	<b>99.63</b>	<b>99.71</b>	<b>99.63</b>	<b>99.84</b>

The combination of ROI-specific feature learning and residual networks allows our model to outperform even advanced approaches like [95], improving accuracy by **1.77%** and maintaining consistency across all statistical measures. Thus, in sync with the overall research inferences, it can be stated that the use of ROI-specific deep-spatio-textural features can make diabetic retinopathy analysis more accurate, even for the multi-class DR classification or allied CAD demands. In comparison to the other state-of-arts, our proposed model performs scalable outputs with high reliability and efficiency

## 4.6 Discussion and Key Insights

From the results in Sections 4.2 through 4.5, we observe the following:

- **FFCM-driven segmentation** consistently isolates lesion regions with higher precision than traditional methods, especially on microaneurysms which are otherwise hard to distinguish from background noise.
- **Residual networks (ResNet50)** significantly outperform classical CNNs (e.g., AlexNet), particularly when combined with handcrafted features. This is attributed to ResNet's skip connections which counteract gradient vanishing in small ROI patches.
- The **hybrid descriptor**: merging GLCM with ResNet features delivers state-of-the-art performance across lesion types, even on datasets not explicitly trained on.
- In terms of **generalizability**, the proposed pipeline maintains high performance across all datasets and abnormalities. Unlike most existing systems that focus on binary or single-lesion detection, our model is scalable for **multi-class DR classification**.

## 5. Conclusion and Future Work

This study presents a novel multi-class diabetic retinopathy (DR) prediction model that integrates advanced segmentation, feature extraction, and classification techniques to deliver a scalable and cost-effective computer-aided diagnosis (CAD) system. Unlike traditional single-lesion approaches, the proposed pipeline handles multiple DR abnormalities: exudates, microaneurysms, and hemorrhages within a unified, real-time-capable framework. The model incorporates cue-independent region of interest (ROI) segmentation via a Firefly-enhanced fuzzy c-means (FFCM) algorithm, allowing precise localization of lesions with minimal post-processing. Preprocessing stages such as histogram equalization, intensity normalization, and Z-score standardization ensure consistency across diverse fundus image datasets.

For feature extraction, the model fuses handcrafted spatio-textural descriptors (GLCM) with high-dimensional deep embeddings from **ResNet50**, leveraging the strengths of residual learning. The combined features are passed to a Random Forest classifier, which efficiently distinguishes between normal and abnormal cases, and further categorizes abnormalities into exudates, microaneurysms, or hemorrhages.

On benchmark datasets, the model achieves an average accuracy of 99.77%, precision of 99.88%, recall of 99.64%, F1-score of 99.75%, sensitivity of 99.64%, and specificity of 99.86%, significantly outperforming state-of-the-art methods across all lesion types.

We are currently collaborating with board-certified ophthalmologists in Bangalore, India, to conduct a blinded, independent grading study for clinical validation of the model. This will be a critical step toward real-world deployment.

In conclusion, the proposed DR-CAD model delivers high diagnostic accuracy, robustness, and real-time capability, offering a valuable clinical decision support tool for early DR screening and severity analysis.

---

## References

- [1] S. Vujosevic, C. Limoli, L. Luzi, et al., "Digital Innovations for retinal care in diabetic retinopathy, *Acta Diabeto*, 2022, Vol. 50, pp. 1521-1530, 2022.
- [2] N. Ahmad, K. -T. Lai and M. Tanveer, "Retinal Blood Vessel Tracking and Diameter Estimation via Gaussian Process With Rider Optimization Algorithm," in *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 3, pp. 1173-1184, March 2024.
- [3] V. I. Agughasi, "The Superiority of Fine-tuning over Full-training for the Efficient Diagnosis of COPD from CXR Images," *Intel. Artif.*, vol. 27, no. 74, pp. 62–79, May 2024, doi: 10.4114/intartif.vol27iss74pp62-79.
- [4] Z. Wang, S. Chen, T. Liu and B. Yao, "Multi-Branching Temporal Convolutional Network With Tensor Data Completion for Diabetic Retinopathy Prediction," in *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 3, pp. 1704-1715, March 2024.
- [5] V. I. Agughasi, "Leveraging Transfer Learning for Efficient Diagnosis of COPD Using CXR Images and Explainable AI Techniques," *Intel. Artif.*, vol. 27, no. 74, Art. no. 74, Jun. 2024, doi: 10.4114/intartif.vol27iss74pp133-151.
- [6] C.-X. Huang, "Long-term effects of pattern scan laser pan-retinal photocoagulation on diabetic retinopathy in chinese patients: A retrospective study," *Int. J. Ophthalmol.*, vol. 13, no. 2, pp. 239\_245, Feb. 2020.
- [7] A. Jabbar et al., "A Lesion-Based Diabetic Retinopathy Detection Through Hybrid Deep Learning Model," in *IEEE Access*, vol. 12, pp. 40019-40036, 2024.
- [8] M. Mateen, J. Wen, M. Hassan, N. Nasrullah, S. Sun, and S. Hayat, "Automatic Detection of Diabetic Retinopathy: A Review on Datasets, Methods and Evaluation Metrics", *IEEE Access*, Vol. 8, 2020, pp. 48784-48811.
- [9] C. Kou, W. Li, Z. Yu, and L. Yuan, "An Enhanced Residual U-Net for Microaneurysms and Exudates Segmentation in Fundus Images", *IEEE Access*, 2020, Vol. 8, pp. 185514- 185525.
- [10] J. Wang, Y. Bai, and B. Xia, "Feasibility of Diagnosing Both Severity and Features of Diabetic Retinopathy in Fundus Photography", *IEEE Access*, 2019, Vol. 7, pp. 102589- 102597.
- [11] H. Leopold, J. Orchard, J. Zelek, and V. Lakshminarayanan, "PixelBNN: Augmenting the PixelCNN with batch normalization and the presentation of a fast architecture for retinal vessel segmentation," *J. Imag.*, vol. 5(2), p. 26, 2019.
- [12] D. Mahapatra, B. Bozorgtabar, and R. Garnavi, "Image super-resolution using progressive generative adversarial networks for medical image analysis," *Comput. Med. Imag. Graph.*, vol. 71, pp. 30\_39, Jan. 2019.
- [13] X. Wang, X. Jiang, and J. Ren, "Blood vessel segmentation from fundus image by a cascade classification framework," *Pattern Recognit.*, vol. 88, pp. 331\_341, 2019.
- [14] Z. Fan, J. Lu, C. Wei, H. Huang, X. Cai, and X. Chen, "A hierarchical image matting model for blood vessel segmentation in fundus images," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2367\_2377, May 2019.
- [15] N. I. Hossain and S. Reza, "Blood vessel detection from fundus image using Markov random \_eld based image segmentation," in *Proc. 4th Int. Conf. Adv. Electr. Eng. (ICAEE)*, Sep. 2017, pp. 123\_127.
- [16] A. M. R. R. Bandara and P. W. G. R. M. P. B. Giragama, "A retinal image enhancement technique for blood vessel segmentation algorithm," in *Proc. IEEE Int. Conf. Ind. Inf. Syst. (ICIIS)*, Dec. 2017, pp. 1\_5.
- [17] K. M. Adal, P. G. van Etten, J. P. Martinez, K.W. Rouwen, K. A. Vermeer, and L. J. van Vliet, "An automated system for the detection and classification of retinal changes due to red lesions in longitudinal fundus images," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 6, pp. 1382\_1390, Jun. 2018.
- [18] P. Costa, A. Galdran, M. I. Meyer, M. Niemeijer, M. Abramoff, A. M. Mendonca, and A. Campilho, "End-to-end adversarial retinal image synthesis," *IEEE Trans. Med. Imag.*, vol. 37, no. 3, pp. 781\_791, Mar. 2018.
- [19] A. V. Ikechukwu and S. Murali, "xAI: An Explainable AI Model for the Diagnosis of COPD from CXR Images," in *2023 IEEE 2nd International Conference on Data, Decision and Systems (ICDDS)*, IEEE, 2023, pp. 1–6. Accessed: Jan. 05, 2025. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10434619/>.
- [20] M. B. Patwari, R. R. Manza, Y. M. Rajput, D. D. Rathod, M. Saswade, and N. Deshpande, "Classification and calculation of retinal blood vessels parameters," in *Proc. IEEE Int. Conf. Converg. Technol.*, Pune, India, Jul. 2016, pp. 1\_6.
-

- [21] V. I. Agughasi and M. Srinivasiah, "Semi-supervised labelling of chest x-ray images using unsupervised clustering for ground-truth generation," *Appl. Eng. Technol.*, vol. 2, no. 3, Art. no. 3, Sep. 2023, doi: 10.31763/aet.v2i3.1143..
- [22] J. Odstrcilik, R. Kolar, T. Kubena, P. Cernosek, A. Budai, J. Hornegger, J. Gazarek, O. Svoboda, J. Jan, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: Evaluation on a new high-resolution fundus image database," *IET Image Proc.* vol. 7(4), pp. 373\_383, 2013.
- [23] D. Relan, T. MacGillivray, L. Ballerini, and E. Trucco, "Retinal vessel classification: Sorting arteries and veins," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 7396\_7399.
- [24] D. Calvo, M. Ortega, M. G. Penedo, and J. Rouco, "Automatic detection and characterisation of retinal vessel tree bifurcations and crossovers in eye fundus images," *Comput. Methods Programs Biomed.*, vol. 103, no. 1, pp. 28\_38, 2011.
- [25] C. Muramatsu, Y. Hatanaka, T. Iwase, T. Hara, and H. Fujita, "Automated detection and classification of major retinal vessels for determination of diameter ratio of arteries and veins," *Proc. SPIE*, vol. 7624, Mar. 2010, Art. no. 76240J.
- [26] J. Paulus, J. Meier, R. Bock, J. Hornegger, and G. Michelson, "Automated quality assessment of retinal fundus photos," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 5, no. 6, pp. 557\_564, Nov. 2010.
- [27] F. M. Villalobos-Castaldi, E. M. Felipe-Riverón, and L. P. Sánchez-Fernández, "A fast, efficient and automated method to extract vessels from fundus images," *J. Visualizat.*, vol. 13, no. 3, pp. 263\_270, Aug. 2010.
- [28] D. S. W. Ting, L. R. Pasquale, L. Peng, J. P. Campbell, A. Y. Lee, R. Raman, G. S. W. Tan, L. Schmetterer, P. A. Keane, and T. Y. Wong, "Artificial intelligence and deep learning in ophthalmology," *Brit. J. Ophthalmol.*, vol. 103, pp. 167\_175, Feb. 2019.
- [29] S. Moccia, E. De Momi, S. El Hadji, and L. S. Mattos, "Blood vessel segmentation algorithms: Review of methods, datasets and evaluation metrics," *Comput. Methods Programs Biomed.*, vol. 158, pp. 71\_91, 2018.
- [30] J. Almotiri, K. Elleithy, and A. Elleithy, "Retinal vessels segmentation techniques and algorithms: A survey," *Appl. Sci.*, vol. 8, no. 2, p. 155, 2018.
- [31] A. Singh, M. K. Dutta, M. ParthaSarathi, V. Uher, and R. Burget, "Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image," *Comput. Methods Programs Biomed.*, vol. 124, pp. 108\_120, Feb. 2016.
- [32] I. Agughasi Victor and S. Murali, "i-Net: a deep CNN model for white blood cancer segmentation and classification," *Int. J. Adv. Technol. Eng. Explor.*, vol. 9, no. 95, Oct. 2022, doi: 10.19101/IJATEE.2021.875564.
- [33] A. R. Chowdhury, T. Chatterjee, and S. Banerjee, "A random forest classifier-based approach in the detection of abnormalities in the retina," *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 193\_203, Jan. 2019.
- [34] C. Lam, C. Yu, L. Huang, and D. Rubin, "Retinal lesion detection with deep learning using image patches," *Investigative Ophthalmol. Vis. Sci.*, vol. 59, no. 1, pp. 590\_596, Jan. 2018.
- [35] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning-based approach for red lesion detection in fundus images," *Comp. Methods Prog. Biomed.*, vol. 153, pp. 115\_127, Jan. 2018.
- [36] W. Cao, N. Czarnek, J. Shan, and L. Li, "Microaneurysm detection using principal component analysis and machine learning methods," *IEEE Trans. Nanobiosci.*, vol. 17, no. 3, pp. 191\_198, Jul. 2018.
- [37] R. Srivastava, D. W. K. Wong, L. Duan, J. Liu, and T. Y. Wong, "Retinal lesion detection in retinal fundus images using frangi-based filters," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2015, pp. 5663\_5666.
- [38] Y. Hatanaka, T. Inoue, S. Okumura, C. Muramatsu, and H. Fujita, "Automated microaneurysm detection method based on double-ring filter and feature analysis in retinal fundus images," in *Proc. 25th IEEE Int. Symp. Comput.-Based Med. Syst.*, Jun. 2012, pp. 1\_4.
- [39] R. Venkatesan, P. Chandakkar, B. Li, and H. K. Li, "Classification of diabetic retinopathy images using multi-class multiple-instance 40 based on color correlogram features," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 1462\_1465.
- [47] S. G. Karst, M. Salas, J. Hafner, C. Scholda, W.-D. Vogl, W. Drexler, M. Pircher, and U. Schmidt-Erfurth, "Three-dimensional analysis of retinal microaneurysms with adaptive optics optical coherence tomography," *Retina*, vol. 39, no. 3, pp. 465\_472, 2019.
- [48] A. Shah, S. Lynch, M. Niemeijer, R. Amelon, W. Clarida, J. Folk, S. Russell, X. Wu, and M. D. Abramoff, "Susceptibility to misdiagnosis of adversarial images by deep learning based retinal image analysis algorithms," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 1454\_1457.
-

- [49] T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus," *Comput. Biomed. Res.*, vol. 29, no. 4, pp. 284\_302, 1996, doi: 10.1006/cbmr.1996.0021.
- [50] T. Walter, J. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1236\_1243, Oct. 2002.
- [51] A. Sopharak, B. Uyyanonvara, S. Barman, and T. H. Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods," *Computer-ized Med. Imag. Graph.*, vol. 32, no. 8, pp. 720\_727, Dec. 2008.
- [52] A. Victor Ikechukwu, P. Sreyas, A. Sena, H. Preetham, and K. Raksha, "Explainable Deep Learning Model for Covid-19 Diagnosis," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 04, no. 07, pp. 3051-3059, Jul. 2022.
- [53] M. M. Fraz, W. Jahangir, S. Zahid, M. M. Hamayun, and S. A. Barman, "Multiscale segmentation of exudates in retinal images using contextual cues and ensemble classification," *Biomed. Signal Process. Control*, vol. 35, pp. 50\_62, May 2017.
- [54] X. Zhang, "Exudate detection in color retinal images for mass screening of diabetic retinopathy," *Med. Image Anal.*, vol. 18, no. 7, pp. 1026\_1043, 2014.
- [55] A. Mizutani, C. Muramatsu, Y. Hatanaka, S. Suemori, T. Hara, and H. Fujita, "Automated microaneurysm detection method based on double ring filter in retinal fundus images," *Proc. SPIE*, vol. 7260, Mar. 2009, Art. no. 72601N.
- [56] G. Quelled, M. Lamard, P. M. Josselin, G. Cazuguel, B. Cochener, and C. Roux, "Optimal wavelet transform for the detection of microaneurysms in retina photographs," *IEEE Trans. Med. Imag.*, vol. 27, no. 9, pp. 1230\_1241, Sep. 2008.
- [57] L. Giancardo, F. Meriaudeau, T. P. Karnowski, Y. Li, S. Garg, K. W. Tobin, and E. Chaum, "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," *Med. Image Anal.*, vol. 16, no. 1, pp. 216\_226, Jan. 2012.
- [58] B. Harangi, B. Antal, and A. Hajdu, "Automatic exudate detection with improved Naïve-Bayes classifier," in *Proc. IEEE Conf.*, Jun. 2012, pp. 1\_4.
- [59] R. Jalloul, C. H. Krishnappa, V. I. Agughasi, and R. Alkhatib, "Enhancing Early Breast Cancer Detection with Infrared Thermography: A Comparative Evaluation of Deep Learning and Machine Learning Models," *Technologies*, vol. 13, no. 1, Art. no. 1, Jan. 2025, doi: 10.3390/technologies13010007.
- [60] P. Chudzik, S. Majumdar, F. Calivá, B. Al-Diri, and A. Hunter, "Microaneurysm detection using fully convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 158, pp. 185\_192, May 2018.
- [61] C. Kou, W. Li, Z. Yu, and J. Hao, "Microaneurysms segmentation with a U-Net based on recurrent residual convolutional neural network," *J. Med. Imag.*, vol. 6, no. 2, p. 025008, Accessed on: May 2019.
- [62] O. Perdomo, J. Arevalo, and F. A. Gonzalez, "Convolutional network to detect exudates in eye fundus images of diabetic subjects," in *Proc. SIPAIM*, vol. 10160, Oct. 2017, Art. no. 101600T.
- [63] S. Bhimshetty and A. V. Ikechukwu, "Energy-efficient deep Q-network: reinforcement learning for efficient routing protocol in wireless internet of things," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 33, no. 2, Art. no. 2, Feb. 2024, doi: 10.11591/ijeecs.v33.i2.pp971-980.
- [64] J. H. Tan, H. Fujita, S. Sivaprasad, S. V. Bhandary, A. K. Rao, K. C. Chua, and U. R. Acharya, "Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network," *Inf. Sci.*, vol. 420, pp. 66\_67, Dec. 2017.
- [65] Z. Feng, J. Yang, L. Yao, Y. Qiao, Q. Yu, and X. Xu, "Deep Retinal Image Segmentation: A FCN-Based Architecture with Short and Long Skip Connections for Retinal Image Segmentation," in *Proc. ICONIP*, 2017, pp. 713\_722.
- [66] R. Zheng, "Detection of exudates in fundus photographs with imbalanced learning using conditional generative adversarial network," *Biomed. Opt. Express*, vol. 9, no. 10, pp. 4863\_4878, Oct. 2018.
- [67] J. Wu, S. Zhang, Z. Xiao, F. Zhang, L. Geng, S. Lou, and M. Liu, "Hemorrhage detection in fundus image based on 2D Gaussian fitting and human visual characteristics," *Opt. Laser Tech.*, vol. 110, pp. 69\_77, 2019.
- [68] H. Zhao, H. Li, S. Maurer-Stroh, Y. Guo, Q. Deng, and L. Cheng, "Supervised segmentation of un-annotated retinal fundus images by synthesis," *IEEE Trans. Med. Imag.*, vol. 38, no. 1, pp. 46\_56, Jan. 2019.
-

- [69] L. Tang, M. Niemeijer, J. M. Reinhardt, M. K. Garvin, and M. D. Abramoff, "Splat feature classification with application to retinal hemorrhage detection in fundus images," *IEEE Trans. Med. Imag.*, vol. 32, no. 2, pp. 364\_375, 2014.
- [70] L. Tang, M. Niemeijer, and M. D. Abramoff, "Splat feature classification: Detection of the presence of large retinal hemorrhages," in *Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro*, Mar. 2011, pp. 681\_684.
- [71] G. B. Kande, T. S. Savithri, and P. V. Subbaiah, "Automatic detection of microaneurysms and hemorrhages in digital fundus images," *J. Digit. Imag.*, vol. 23, no. 4, pp. 430\_437, Aug. 2010.
- [72] P. Prenta<sup>2</sup>iç and S. Lonçariç, "Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion," *Comput. Methods Programs Biomed.*, vol. 137, pp. 281\_292, Dec. 2016.
- [73] P. Prentasic and S. Loncaric, "Detection of exudates in fundus photographs using convolutional neural networks," in *Proc. 9th Int. Symp. Image Signal Process. Anal. (ISPA)*, Sep. 2015, pp. 188\_192.
- [74] M. Omar, F. Kheli\_, and M. A. Tahir, "Detection and classification of retinal fundus images exudates using region based multiscale LBP texture approach," in *Proc. Int. Conf. Control, Decis. Inf. Technol*, Apr. 2016, pp. 227\_232.
- [75] B. Harangi, I. Lazar, and A. Hajdu, "Automatic exudate detection using active contour model and regionwise classification," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 5951\_5954.
- [76] A. Osareh, B. Shadgar, and R. Markham, "A computational-intelligence based approach for detection of exudates in diabetic retinopathy images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 4, pp. 535\_545, Jul. 2009.
- [77] A. Sopharak, U. Bunyarit, and S. Barman, "Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy C-means clustering," *Sensors*, vol. 9, no. 3, pp. 2148\_2161, 2009.
- [78] D. Shruthi, H. K. Chethan, and V. I. Agughasi, "Effective Approach for Fine-Tuning Pre-Trained Models for the Extraction of Texts From Source Codes," in *ITM Web of Conferences*, EDP Sciences, 2024, p. 03004. Accessed: Jan. 05, 2025. [Online]. Available: [https://www.itm-conferences.org/articles/itmconf/abs/2024/08/itmconf\\_icmaetm2024\\_03004/itmconf\\_icmaetm2024\\_03004.html](https://www.itm-conferences.org/articles/itmconf/abs/2024/08/itmconf_icmaetm2024_03004/itmconf_icmaetm2024_03004.html).
- [79] B. Harangi and A. Hajdu, "Automatic exudate detection by fusing multiple active contours and regionwise classification," *Comput. Biol. Med.*, vol. 54, pp. 156\_171, Nov. 2014.
- [80] M. García, C. I. Sánchez, M. I. López, D. Abásolo, and R. Hornero, "Neural network-based detection of hard exudates in retinal images," *Comput. Methods Programs Biomed.*, vol. 93, no. 1, pp. 9\_19, 2009.
- [81] F. Zabihollahy, A. Lochbihler, and E. Ukwatta, "Deep learning-based approach for fully automated detection and segmentation of hard exudate from retinal images," *Proc. SPIE*, vol. 10953, Mar. 2019, Art. no. 1095308, doi: 10.1117/12.2513034.
- [82] Z. Yan, X. Han, C. Wang, Y. Qiu, Z. Xiong, and S. Cui, "Learning mutually local-global U-Nets for high-resolution retinal lesion segmentation in fundus images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Venice, Italy, Apr. 2019, pp. 597\_600.
- [83] V. I. Agughasi, Y. Dk, and D. M. Snehil, "Early Prognosis of Heart Failure from Clinical Symptoms using K-Means and Naïve Bayes Algorithms," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 9, no. 7, 2020, doi: DOI 10.17148/IJARCCCE.2020.9709.
- [84] R. Casanova, S. Saldana, E. Y. Chew, R. P. Danis, C. M. Greven, and W. T. Ambrosius, "Application of random forests methods to diabetic retinopathy classification analyses," *PLoS One*, vol. 9, no. 6, p. e98587, 2014.
- [85] A. Kwasigroch, B. Jarzembinski, and M. Grochowski, "Deep CNN based decision support system for detection and assessing the stage of diabetic retinopathy," in *Proc. Int. Interdiscipl. PhD Workshop*, 2018, pp. 111\_116.
- [86] J. Shan and L. Li, "A deep learning method for microaneurysm detection in fundus images," in *Proc. IEEE 1st Int. Conf. Connected Health, Appl., Syst. Eng. Technol. (CHASE)*, Jun. 2016, pp. 357\_358.
- [87] M. Mateen, J. Wen, Nasrullah, S. Song, and Z. Huang, "Fundus image classification using VGG-19 architecture with PCA and SVD," *Symmetry*, vol. 11, no. 1, p. 1, 2019.
- [88] L. Wu, C. Wan, Y. Wu, and J. Liu, "Generative caption for diabetic retinopathy images," in *Proc. Int. Conf. Secur., Pattern Anal., Cybern. (SPAC)*, Dec. 2017, pp. 515\_519.
- [89] R. Tennakoon, D. Mahapatra, P. Roy, S. Sedai, and R. Garnavi, "Image quality classification for DR screening using convolutional neural networks," in *Proc. Ophthalmic Med. Image Anal. 3rd Int. Workshop*, Oct. 2016, pp. 1\_9.
- [90] M. Purandare and K. Noronha, "Hybrid system for automatic classification of diabetic retinopathy using fundus images," in *Proc. Online Int. Conf. Green Eng. Technol. (IC-GET)*, Nov. 2016, pp. 1\_5.
- [91] D. K. Prasad, L. Vibha, and K. R. Venugopal, "Early detection of diabetic retinopathy from digital retinal fundus images," in *Proc. IEEE Recent Adv. Intell. Comput. Syst. (RAICS)*, Dec. 2015, pp. 240\_245.
-

[92] K. Adem, M. Hekim, and S. Demir, "Detection of hemorrhage in retinal images using linear classifiers and iterative thresholding approaches based on Firefly and particle swarm optimization algorithms," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 27, no. 1, pp. 499\_515, 2019.

[93] A. Victor Ikechukwu and M. S, "CX-Net: an efficient ensemble semantic deep neural network for ROI identification from chest-x-ray images for COPD diagnosis," *Mach. Learn. Sci. Technol.*, vol. 4, no. 2, p. 025021, Jun. 2023, doi: 10.1088/2632-2153/acd2a5.

[94] L. Guo, J.-J. Yang, L. Peng, J. Li, and Q. Liang, "A computer-aided healthcare system for cataract classification and grading based on fundus image analysis," *Comput. Ind.*, vol. 69, pp. 72\_80, May 2015.

[95] P. Khojasteh, L. A. P. Júnior, T. Carvalho, E. Rezende, B. Aliahmad, J. P. Papa, and D. K. Kumar, "Exudate detection in fundus images using deeply-learnable features," *Comput. Biol. Med.*, vol. 104, pp. 62\_69, Jan. 2019.

[96] J. Kaur and D. Mittal, "A generalized method for the segmentation of exudates from pathological retinal fundus images," *Biocybernetics Biomed. Eng.*, vol. 38, no. 1, pp. 27\_53, 2018.

[97] A. V. Ikechukwu, M. S, and H. B, "COPDNet: An Explainable ResNet50 Model for the Diagnosis of COPD from CXR Images," in *2023 IEEE 4th Annual Flagship India Council International Subsections Conference (INDISCON)*, Mysore, India: IEEE, Aug. 2023, pp. 1–7. doi: 10.1109/INDISCON58499.2023.10270604.

[98] A. M. Alhassan, "Driving Training-Based Optimization- Multitask Fuzzy C-Means (DTBO-MFCM) Image Segmentation and Robust Deep Learning Algorithm for Multi-center Breast Histopathological Images," in *IEEE Access*, vol. 11, pp. 136350-136360, 2023.

[99] A. Bharathi, S. R. Kumar, "Evolutionary computing Firefly Algorithm Assisted Fuzzy C-Means Clustering Model for Steel Strip Surface Crack Detection", *International Journal of Advanced Science and Technologies*, Vol. 29 (08), pp. 6118-6138, 2020.

[100] S. Prathibha and Siddapaji, "Advancing diabetic retinopathy diagnosis with fundus imaging: A comprehensive survey of computer-aided detection, grading and classification methods", *ScienceDirect, Global Transitions (KEAD)*, Vol. 6, pp. 93-112, 2024.

[101] B. Keerthiveena, T. Veerakumar, S. Esakkirajan and B. N. Subudhi, "Computer-aided diagnosis for Diabetic Retinopathy based on Firefly algorithm," *2019 11th International Conference on Advanced Computing (ICoAC)*, Chennai, India, 2019, pp. 310-315.

[102] P. M. Ebin, N. M. Mathkunti and U. Ananthanagu, "Revolutionizing Diabetic Retinopathy Diagnosis:Harnessing Automated Deep Learning Classifiers for Early Detection," *2023 Global Conference on Information Technologies and Communications (GCITC)*, Bangalore, India, 2023, pp. 1-6.

[103] S. Akhtar, S. Aftab, M. Ahmad and B. Ihnaini, "A Transfer Learning Based Framework for Diabetic Retinopathy Detection Using Data Fusion," *2024 2nd International Conference on Cyber Resilience (ICCR)*, Dubai, United Arab Emirates, 2024, pp. 1-5.

[104] R. Patil, A. Bhamare, S. S. Devdas, A. Bandawar, A. Ahirkar and S. Chaudhari, "Expeditious Detection of Diabetic Retinopathy Using ResNet 50 CNN Algorithm," *2023 3rd International Conference on Advancement in Electronics & Communication Engineering (AECE)*, GHAZIABAD, India, 2023, pp. 224-226.

[105] S. Lakhera and A. Garg, "Classification of Diabetic Retinopathy Using Slime Mould Optimization Based ResNet-18 Deep Learning Model," *2024 International Conference on Smart Systems for applications in Electrical Sciences (ICSSSES)*, Tumakuru, India, 2024.

## APPENDICES

### Appendix A: Bayesian-Optimised Hyper-parameter Trials

Trial ID	Learning-rate	Drop-out	Random-Forest Trees	5-fold Val. Accuracy (%)
T-01	$3 \times 10^{-4}$	0.30	80	97.9
T-02★	$1 \times 10^{-4}$	<b>0.50</b>	<b>60</b>	<b>98.4</b>
T-03	$5 \times 10^{-5}$	0.50	60	98.1
T-04	$1 \times 10^{-4}$	0.40	80	98.0
T-05	$3 \times 10^{-4}$	0.60	60	97.5

★ Configuration reported in the main manuscript.

### Appendix B: Additional Segmentation Benchmarks

Segmentation Model	Parameters (M)	Dice Coefficient	Matthews Corr. Coeff. (MCC)	Imbalance Strategy
Firefly-FCM (proposed)	<b>2.1</b>	<b>0.943</b>	<b>0.962</b>	Adaptive ROI + ensemble voting
<b>Residual U-Net</b>	7.5	0.928	0.944	Dice-loss weighting
<b>Vanilla U-Net</b>	7.0	0.901	0.931	None

### Appendix C: Implementation & Runtime Environment

Category	Specification
<b>GPU</b>	NVIDIA GTX1060 (6 GB)
<b>CPU</b>	Intel Core i7, 8 <sup>th</sup> Gen
<b>RAM</b>	24 GB DDR4-3200
<b>OS</b>	Ubuntu 22.04 LTS
<b>Key Libraries</b>	PyTorch 2.2, Torchvision 0.17, OpenCV 4.10, scikit-learn 1.5
<b>Training Time</b>	3 h 17 m for 50 epochs (segmentation) + 2 h 02 m for 40 epochs (classification)
<b>Total Parameters</b>	Segmentation $\approx$ 2.1 M • Classification head $\approx$ 24.4 M

### Appendix D: Dataset References

Hoover A., Kouznetsova V., Goldbaum M., “Locating Blood Vessels in Retinal Images by Piecewise Threshold Probing of a Matched Filter Response,” *IEEE Transactions on Medical Imaging*, 2000. (STARE)

Kauppi T. et al., “The DIARETDB1 Diabetic Retinopathy Database and Evaluation Protocol,” *British Machine Vision Conference (BMVC)*, 2007. (DIARET-DB1)

Niemeijer M. et al., “Retinopathy Online Challenge,” *IEEE Transactions on Medical Imaging*, 2010. (DIRECT-DB1)

DOK-MCM Dataset, introduced in Thai diabetic retinopathy screening research at King Chulalongkorn Memorial Hospital. (DOK-MCM)

Decenciere E. et al., “Feedback on a Publicly Distributed Image Database: the Messidor Database,” *Image Analysis & Stereology*, 2014. (Messidor-1)