Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 4, October 2026, pp: 1314-1330 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

A Dual-Stage Deep Lesion Segmentation Framework with Gradient and Boundary Optimization for Diabetic Retinopathy Retinal Images

Malathi P[1], Dhinakaran D[2], Jeyalakshmi S[3], Kavitha P[4], Palpandi S[2] and Prabaharan G[2]

[1] Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, India
[2] Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India
[3] Department of Artificial Intelligence and Data Science, P.S.R. Engineering College, Sivakasi, India
[4] Department of Computer Science and Engineering, R.M.K Engineering College, Chennai, India

Corresponding author: Dhinakaran D. (e-mail: drdhinakarand@veltech.edu.in), **Author(s) Email**: Malathi P (e-mail: malathip@srmist.edu.in), Jeyalakshmi S (e-mail: jeyalakshmisamyraj46@gmail.com), Kavitha P (e-mail: pka.cse@rmkec.ac.in), Paul Pandi P (e-mail: sppecerdeee@gmail.com), Prabaharan G (e-mail: drprabaharang@veltech.edu.in)

Abstract **Diabetic Retinopathy (DR) is a leading global cause of preventable blindness, and the early detection and precise segmentation of pathologies play a key role in clinical decisions in this disease. Conventional deep learning models, however, suffer from the problem of weak lesion boundaries, gradient inconsistency, and structural distortion in heterogeneous datasets. To overcome these drawbacks, we present a sophisticated systemthat combines GEONet (Gradient and Edge-Optimized Network) to achieve accurate gradient- and edge-aware segmentation, and BESNet (Boundary-Enhanced Segmentation Network) to introduce boundary refinement and structural preservation. This two-step strategy can guarantee the high-fidelity capture of subtle retinal lesions with anatomical consistency. Comparative analysis was carried out against state-of-the-art baselines. Experimental results on the EyePACS dataset demonstrate the effectiveness of the proposed framework. The integrated GEONet+BESNet architecture achieved a Dice Coefficient of 89.0%, Intersection-over-Union (IoU) of 87.0%, Boundary Accuracy of 89.0%, and a Hausdorff Distance of 5.1, outperforming all competing methods in terms of segmentation fidelity and boundary preservation. Structural consistency was further enhanced, attaining a Dice Similarity Coefficient (DSC) of 90.5% and a Structural Similarity Index Measure (SSIM) of 90.3%. From a clinical screening perspective, the proposed framework achieved a Precision of 92.8% and Specificity of 94.6%, indicating reliable lesion localization with a reduced rate of false-positive detections. These findings confirm that the synergistic integration of gradient-aware segmentation through GEONet and boundary-enhanced refinement through BESNet effectively preserves lesion morphology and improves segmentation robustness across heterogeneous retinal imaging conditions.**

Keywords **Diabetic Retinopathy; Lesion Boundaries; Gradient and Edge–Optimized Network; Boundary–Enhanced Segmentation Network; Deep Learning Models; Morphology; Edge Optimization**.

I. Introduction

Diabetic retinopathy (DR) is a major diabetes-related complication and a leading cause of preventable vision loss worldwide, particularly among working-age populations [1]. Early identification of retinal lesions such as microaneurysms, hemorrhages, and exudates is essential for preventing disease progression. Automated retinal image analysis has therefore emerged as an important tool for supporting large-scale screening and reducing clinical workload. However, accurate retinal lesion segmentation remains challenging due to variations in image quality, illumination, contrast, lesion morphology, and class imbalance across datasets [2-4]. In addition, precise delineation of lesion boundaries is often difficult, reducing segmentation reliability. Representative DR

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

images spanning different severity levels [5]. These challenges highlight the need for robust segmentation frameworks that preserve lesion structures and maintain consistent performance across heterogeneous retinal imaging datasets [6,7].

Consequently, there is an urgent need to develop automated, accurate, and scalable diagnostic frameworks that can supplement or even substitute clinical workflows in both urban and rural environments [8-10]. Motivated by these pressing challenges, the present work identifies several core motivations and translates them into clear objectives that guide the design of the proposed framework. This study presents the GEONet–BESNet framework for retinal lesion segmentation in diabetic retinopathy images. GEONet enhances gradient and edge preservation for accurate lesion localization, while BESNet improves boundary fidelity and structural consistency. Together, they provide robust and accurate segmentation across diverse datasets, supporting reliable retinal image analysis under heterogeneous imaging conditions. The remainder of this paper is organized as follows. Section 2 reviews existing diabetic retinopathy segmentation and analysis methods and identifies key research gaps. Section 3 presents the proposed GEONet–BESNet framework. Section 4 describes the experimental setup, datasets, evaluation metrics, and comparative results. Finally, Section 5 concludes the study, discusses its limitations, and outlines future research directions for improving retinal lesion segmentation and clinical applicability.

II. Related Works

Diabetic retinopathy (DR) analysis remains challenging due to variations in image quality, illumination, lesion appearance, and class imbalance. Although deep learning methods have improved lesion detection and segmentation, many suffer from poor boundary preservation, structural inconsistency, and limited cross-dataset generalization. These limitations highlight the need for robust edge-aware segmentation frameworks capable of accurately delineating retinal lesions while maintaining reliable performance across heterogeneous retinal imaging datasets.

Recent advances in diabetic retinopathy (DR) analysis have focused on improving lesion detection, segmentation, and grading through deep learning and attention-based architectures. Mukherjee et al. [11] proposed BhAFPN+SVM, which integrates multi-scale feature fusion and a hybrid attention mechanism to improve lesion localization and DME grading. Baltov et al. [12] introduced a compound-scaled encoder–decoder framework with scSE attention and compound loss functions to address class imbalance in lesion segmentation. Dai et al. [13] developed DeepDR using over 466,000 retinal images, demonstrating the benefits of large-scale training for lesion detection and grading. Asia et al. [14] investigated preprocessing and regularization strategies with ResNet and VGG architectures, while Malhi et al. [15] utilized lesion-specific analysis based on exudate and microaneurysm characteristics. Praveen et al. [16] combined CLAHE, feature selection, and DeepForest classification, whereas Nazir et al. [17] employed ensemble CNN models to improve robustness. Zedadra et al. [18] extended DR analysis through multimodal fusion of retinal images and clinical attributes. Additional studies explored transfer learning [19], UNet++ and Capsule Networks [20], enhanced deep learning models [21], meta-heuristic optimization [22], denoising-aware feature extraction [23], hybrid deep descriptors [24], and dense feature fusion [25]. Despite these advances, limitations in boundary preservation, structural consistency, computational efficiency, and cross-dataset generalization remain significant challenges.

Existing DR analysis approaches can be categorized into classification, segmentation, attention-based, boundary-aware, and multimodal frameworks. Classification models provide high grading accuracy but limited lesion localization. Segmentation methods directly delineate pathological regions, yet often produce blurred boundaries and inconsistent structural details. Attention-based architectures improve feature discrimination but generally focus on global saliency rather than contour preservation. Boundary-aware approaches enhance edge precision but rarely enforce structural consistency throughout the segmentation process. Multimodal systems improve diagnostic performance by incorporating clinical information; however, their dependence on additional data limits scalability. Consequently, a significant gap remains in the development of a unified framework that simultaneously preserves gradients, maintains lesion boundaries, and ensures robust segmentation across diverse retinal imaging conditions.

To address existing limitations, this study introduces the GEONet–BESNet framework for retinal lesion segmentation. GEONet employs gradient-aware learning to preserve lesion contours and structural details, while BESNet performs boundary-focused refinement to improve segmentation consistency. The framework incorporates morphology-guided preprocessing and class-balancing strategies to enhance robustness against imaging variability and improve sensitivity to subtle lesions. Unlike computationally intensive ensemble or multimodal approaches, the proposed dual-stage architecture achieves accurate lesion delineation while maintaining practical scalability. The integration of gradient optimization and boundary enhancement provides improved structural fidelity, lesion localization, and cross-dataset robustness in retinal image segmentation.

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

III. Method

The proposed GEONet–BESNet framework consists of three stages: preprocessing, GEONet segmentation, and BESNet refinement, as shown in Fig. 1. The preprocessing stage performs normalization, denoising, contrast enhancement, and ROI extraction to improve lesion visibility and generate edge-aware structural cues [26-28]. GEONet then extracts multi-scale features and employs Gradient Streamlined Attention and Edge-Constrained Adversarial Learning to produce accurate lesion segmentation with enhanced boundary awareness. Finally, BESNet refines the coarse segmentation through the Boundary Preservation, Texture-Aware Interior, and Synergistic Fusion modules, thereby generating high-fidelity lesion masks with improved structural consistency, boundary sharpness, and robustness across diverse retinal imaging conditions.

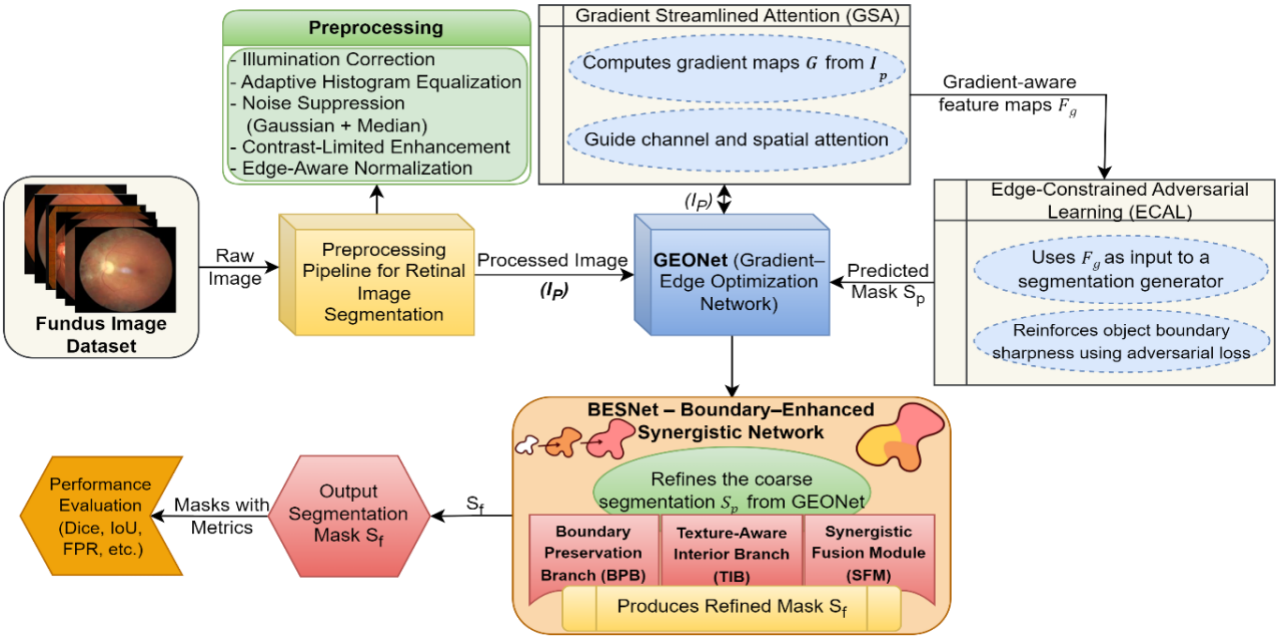

Fig. 1. Dual-Stage Deep Segmentation Framework for Diabetic Retinopathy Screening.

A. Preprocessing Pipeline for Retinal Image Segmentation

To improve segmentation robustness, raw fundus images I_{raw} were transformed into a standardized representation I_{prep} through normalization, illumination correction, contrast enhancement, denoising, and ROI extraction [29-31]. Each color channel was normalized to zero mean and unit variance by using Eq. (1) as follows [4], Each color channel was normalized to zero mean and unit variance using Eq. (1) [4], where $I_{raw}^{c}(x,y)$ denotes the intensity of the input retinal image at pixel (x,y) for color channel $c \in \{R,G,B\}$, $I_n^c(x,y)$ is the normalized pixel intensity, and μ_c and σ_c represent the mean and standard deviation of the corresponding color channel computed exclusively from the training set to ensure dataset invariance and prevent data leakage. Illumination variations were corrected using a Retinex-inspired model by using Eq. (2) as follows [29], where $I_n(x,y)$ denotes the normalized image, $L(x,y)$ represents the slowly varying illumination component, and $R(x,y)$ denotes the reflectance component containing retinal anatomical structures and lesion information.

$$I_n^c(x,y) = \frac{I_{raw}^c(x,y) - \mu_c}{\sigma_c} \quad (1)$$

$$I_n(x,y) = L(x,y) \cdot R(x,y) \quad (2)$$

Taking the logarithm converts the multiplicative model to additive by Eq. (3) as follows [31], where $I_n(x,y)$ is the normalized retinal image, $L(x,y)$ denotes the illumination component, and $R(x,y)$ represents the reflectance component. A Gaussian low-pass filter G_σ is used to estimate $log\, L(x,y)$, and the structural component is obtained by subtraction by using tq. (4) as follows [28], where G_σ denotes a Gaussian kernel with standard deviation σ and * denotes the convolution operator. We employ Contrast-Limited Adaptive Histogram Equalization (CLAHE) by using the Eq. (5) as follows [30], where I_c^G is the illumination-corrected green channel, I_e^G is the enhanced green channel, and L_c denotes the CLAHE clip limit. We apply a bilateral filter, which combines spatial and intensity weighting by using the Eq. (6) as follows [28], where $I_f(x,y)$ is the filtered image, Ω represents the local neighborhood,

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

(i,j) denotes neighboring pixels, W_p is the normalization factor, and σ_s and σ_r control spatial and intensity smoothing, respectively. The filtered image is masked and resized as I_{prep}, as shown in Eq. (7) as follows [4], where M_{ROI} is the binary region-of-interest mask obtained from Hough Circle Transform, H_t and W_t denote the target image height and width, and I_{prep} represents the final preprocessed image used as input to GEONet.

$$log\ I_n(x,y) = log\ L(x,y) + log\ R(x,y) \quad (3)$$

$$log\ R(x,y) = log\ I_n(x,y) - G_\sigma * log\ I_n(x,y) \quad (4)$$

$$I_e^G = H_{CLAHE}(I_c^G, L_c) \quad (5)$$

$$I_f(x,y) = \frac{1}{W_p}\sum_{(i,j)\in\Omega} I_e(i,j) \cdot exp\left(-\frac{\|(i,j)-(x,y)\|^2}{2\sigma_s^2} - \frac{\|I_e(i,j)-I_e(x,y)\|^2}{2\sigma_r^2}\right) \quad (6)$$

The filtered image is masked and resized:

$$I_{prep} = Resize\left(I_f \cdot M_{ROI},\ H_t, W_t\right) \quad (7)$$

In the implementation, CLAHE was applied using a clip limit of 2.0 and a tile grid size of 8 × 8. Illumination correction employed a Gaussian low-pass filter with σ = 15. Bilateral filtering was performed with σ_s = 75 and σ_r = 75 to suppress noise while preserving anatomical edges. ROI extraction was based on the largest retinal field detected using the Hough Circle Transform, and the resulting region was resized to 512 × 512 pixels before model training. Normalization statistics (μc, σc) were computed exclusively from the training set to prevent information leakage during validation and testing.

Table 1. Architectural Distinction of GEONet–BESNet.

✓ = explicitly implemented; ✗ = not included; Partial = indirectly or incompletely addressed in the original method

Feature	Attention U-Net	Boundary Loss Networks	Adversarial Segmentation	Proposed GEONet–BESNet
Spatial Attention	✓	✗	✗	✓
Gradient-guided Attention	✗	✗	✗	✓
Explicit Edge Map Learning	✗	Partial	Partial	✓
Boundary Realism Discriminator	✗	✗	Partial	✓
Multi-scale Boundary Propagation	✗	✗	✗	✓
Texture-aware Interior Refinement	✗	✗	✗	✓
Dual-stage Segmentation Refinement	✗	✗	✗	✓

B. Novelty Positioning Against Existing Segmentation Frameworks

The proposed GEONet–BESNet framework distinguishes itself from existing retinal segmentation methods through the coordinated integration of gradient-guided attention, edge-constrained adversarial learning, and boundary–texture synergistic refinement (Table 1). Unlike Attention U-Net, the Gradient Streamlined Attention (GSA) module incorporates gradient magnitude information to explicitly enhance lesion boundaries and high-frequency pathological structures. Furthermore, the Edge-Constrained Adversarial Learning (ECAL) module enforces boundary realism through adversarial supervision rather than relying solely on boundary-loss optimization. BESNet further refines segmentation using complementary Boundary Preservation and Texture-Aware Interior branches, enabling simultaneous optimization of contour accuracy and interior lesion consistency.

C. GEONet (Gradient–Edge Optimization Network)

GEONet is a hybrid segmentation enhancement framework that integrates fine-grained attention-based feature refinement with precise, boundary-aware adversarial optimization. The architecture fuses Gradient Streamlined Attention (GSA), which selectively amplifies gradient-rich and contextually relevant features, with Edge-Constrained Adversarial Learning (ECAL), which enforces structural and contour fidelity through boundary-aware discriminative supervision. By jointly leveraging GSA's ability to preserve high-frequency details and ECAL's adversarial constraint on shape accuracy, GEONet ensures both region-level consistency and sharp delineation of lesion borders. This dual-component synergy addresses the common limitations of deep segmentation models, which often excel at region localization but falter at maintaining boundary precision in challenging scenarios such as low contrast, irregular morphology, or high noise. Consequently, GEONet serves as an advanced post-augmentation segmentation enhancer, capable of refining coarse segmentation masks into clinically reliable, high-resolution outputs

1. Gradient Streamlined Attention (GSA)

The proposed Gradient Streamlined Attention (GSA) mechanism enhances retinal lesion segmentation by emphasizing gradient-rich, pathology-relevant regions

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

while suppressing redundant background information. GSA combines gradient-prior spatial attention with lightweight channel recalibration to selectively focus on lesion boundaries and high-frequency structural details. By prioritizing informative spatial regions and feature channels, GSA improves lesion localization, boundary preservation, and feature discrimination, enabling more accurate segmentation of subtle pathological structures without introducing significant computational overhead.

a. Spatial Attention Streamlining

Let $F \in R^{C\times H\times W}$ denote an intermediate feature map extracted from a convolutional encoder, where C, H, and W are the number of channels, height, and width, respectively. To introduce spatial selectivity, we first compute a channel-averaged representation as defined in Eq. (8) as follows [5], where F_{avg} denotes the channel-averaged feature map, F_c represents the c^{th} feature channel, and $c = 1, \ldots, C$.

$$F_{avg} = \frac{1}{C}\sum_{c=1}^{C} F_c \quad (8)$$

The spatial gradient magnitude map G_s is then obtained by applying a Sobel operator in the horizontal and vertical directions to F_{avg}, followed by squared summation of the partial derivatives:

$$G_s(i,j) = \left(\frac{\partial F_{avg}}{\partial_x}\right)^2 + \left(\frac{\partial F_{avg}}{\partial_y}\right)^2 \quad (9)$$

This formulation of Eq. (9) as follows [7] serves as a structural saliency detector, emphasizing regions with high gradient responses, such as optic nerve head boundaries or sharp lesion edges, which are diagnostically significant in DR [32]. The gradient magnitude map G_s is then projected into an attention map via a 1×11 convolution followed by a sigmoid activation as defined in Eq. (10) as follows [32]:

$$A_s = \sigma\big(Conv_{1\times1}(G_s)\big) \quad (10)$$

where $\sigma(\cdot)$ denotes the element-wise sigmoid function. The spatially refined feature map is obtained through element-wise multiplication between the original feature map and the attention mask as defined in Eq. (11) as follows [15]

$$F_s = F \otimes A_s \quad (11)$$

where $\otimes$ indicates element-wise multiplication. This operation ensures that activations in low-gradient background areas, such as uniform choroidal regions, are suppressed, while features corresponding to fine vascular branching patterns, hemorrhage borders, or optic disc rims are enhanced.

b. Channel Attention Streamlining

Following spatial refinement, GSA applies a lightweight channel attention mechanism to preserve discriminative feature channels while suppressing redundant or noise-dominated channels. For each channel c, a global descriptor is computed as the average spatial response:

$$z_c = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} F_s(c,i,j) \quad (12)$$

The resulting descriptor vector $z \in R^C$ encodes the global activation strength per channel, capturing disease-specific response patterns such as vessel density fluctuations in DR. Channel attention weights are computed via a two-layer fully connected network with a bottleneck structure to reduce computational overhead. Specifically:

$$A_c = \sigma\big(W_2 \cdot \delta(W_1 \cdot z)\big) \quad (13)$$

The final GSA-refined feature map is obtained by reweighting the spatially refined features F_s with the channel attention map A_c by utilizing Eq. (12-14) as follows [20], where z_c denotes the global descriptor of the c^{th} channel, $F_s(c,i,j)$ represents the spatially refined feature value at channel $c$and pixel location (i,j), and H and W denote the height and width of the feature map, and W_2 are the learnable weight matrices of the first and second fully connected layers, $\delta(\cdot)$ denotes the ReLU activation function, $\sigma(\cdot)$ represents the sigmoid activation function, and A_c is the channel attention vector, denotes the output feature map after Gradient Streamlined Attention, and $\otimes$ represents channel-wise element-wise multiplication, respectively.

$$F_{GSA} = F_s \otimes A_c \quad (14)$$

2. Edge-Constrained Adversarial Learning (ECAL)

The Edge-Constrained Adversarial Learning (ECAL) module preserves fine lesion boundaries by integrating edge-aware constraints with adversarial training. Unlike conventional loss functions that often produce over-smoothed contours, ECAL enforces structural consistency and boundary realism, improving the delineation of diagnostically significant retinal lesions and pathological regions [33-35]. The first step in ECAL involves edge map extraction for both the predicted segmentation S_p and the ground truth S_{gt}. Using a morphological gradient operator, the binary edge map $E(X)$ for a given segmentation mask X is computed as:

$$E(X) = dilate(X) - erode(X) \quad (15)$$

Although the DRD, APTOS, and EyePACS datasets primarily provide image-level diabetic retinopathy severity labels, the proposed segmentation framework requires pixel-level lesion annotations for supervised training and quantitative evaluation. Accordingly, reference lesion masks (S_{gt}) were utilized during model optimization and validation. These masks served as ground-truth annotations for computing the Dice Coefficient, Intersection-over-Union (IoU), Hausdorff Distance, Boundary Accuracy, DSC, and SSIM metrics. Furthermore, the edge maps used in the ECAL module were generated directly from the corresponding ground-truth masks through the morphological gradient

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

operation described in Eq. (15) as follows [34]. Following edge extraction, ECAL employs an adversarial edge loss to enhance the realism of predicted boundaries. A dedicated discriminator D_{edge} is trained to distinguish whether an edge map originates from the ground truth or from the predicted segmentation. The generator's adversarial loss is defined as Eq. (16-19) as follows [35], where L_{adv}denotes the adversarial loss, S_p is the predicted segmentation mask, E_p represents the predicted edge map extracted from S_p, $D_{edge}(\cdot)$ is the edge discriminator, and $\mathbb{E}[\cdot]$ denotes the expectation operator, where S_{gt} denotes the ground-truth segmentation mask and E_{gt} is the corresponding ground-truth edge map.

$$L_{adv} = -E_{S_p}\left[log\ D_{edge}(E_p)\right] \quad (16)$$

which encourages the segmentation network to produce edge maps that the discriminator cannot differentiate from real ones. Correspondingly, the discriminator loss is:

$$L_D = -E_{S_{gt}}\left[log\ D_{edge}(E_{gt})\right] - ES_p\left[log\left(1 - D_{edge}(E_p)\right)\right] \quad (17)$$

This adversarial setting introduces a competitive dynamic between the generator and the discriminator, driving the generator toward improved edge realism in the predicted masks. To further constrain the accuracy of boundaries, a boundary matching loss based on the Dice coefficient is incorporated:

$$L_{edge} = 1 - \frac{2\cdot|E_p \cap E_{gt}|}{|E_p|+|E_{gt}|} \quad (18)$$

The total loss function for ECAL integrates segmentation accuracy, edge alignment, and adversarial feedback:

$$L_{total} = L_{seg} + \lambda_{edge}L_{edge} + \lambda_{adv}L_{adv} \quad (19)$$

where L_{seg} represents a hybrid segmentation loss (combining Dice and Binary Cross-Entropy to balance overlap and pixel-wise accuracy), and λ_{edge} and λ_{adv} are weighting parameters controlling the influence of boundary and adversarial constraints [36-38]. This composite objective ensures that ECAL not only achieves high pixel-wise segmentation accuracy but also preserves diagnostically critical lesion contours, a requirement for robust diabetic retinopathy prediction models.

D. BESNet: Boundary–Enhanced Synergistic Network for Segmentation Refinement

Following the feature extraction and attention-optimization phase through GEONet, the segmentation stage is conducted using the Boundary-Enhanced Synergistic Segmentation Network (BESNet). BESNet is specifically designed to exploit the gradient–edge-optimized outputs from GEONet, ensuring that the final segmentation mask retains high boundary fidelity while achieving a uniform representation of interior lesions. The network operates in a dual-branch refinement architecture: Both the Boundary Preservation Branch (BPB), and Texture-Aware Interior Branch (TIB, are subsequently fused through a Synergistic Fusion Module (SFM).

The objective of BESNet is to produce a refined final segmentation mask $S_f \in R^{H\times W}$ as shown in Eq. (20) as follows [2], where S_f denotes the final refined segmentation mask, S_p is the coarse segmentation mask produced by GEONet, E_p represents the predicted edge map, F_g denotes the gradient-aware feature representation, and $F_{BESNet}(\cdot)$ is the refinement function implemented by BESNet.

$$S_f = F_{BESNet}\left(S_p, E_p, F_g\right) \quad (20)$$

1. Boundary Preservation Branch (BPB)

The BPB is responsible for retaining sharp lesion borders and mitigating blurring that often arises in standard segmentation pipelines. Using S_p and E_p, a Multi-Scale Boundary Propagation (MSBP) process is applied. At each scale $l \in \{1,2,\dots,L\}$, the edge map is dilated and eroded to construct a set of boundary priors B_l, which guide the mask refinement as shown in Eq. (21) as follows [9]. where B_l denotes the boundary prior at the l^{th} scale, L is the total number of propagation scales, r_l is the structuring-element radius at scale l, and $\phi_{dilate}(\cdot, r_l)$ and $\phi_{erode}(\cdot, r_l)$ represent the morphological dilation and erosion operations, respectively.

The boundary refinement update at each scale is then given by Eq. (22) as follows [14], where $S_p^{(l-1)}$ and $S_p^{(l)}$ denote the segmentation masks before and after refinement at the l^{th}scale, respectively, λ_b is the boundary refinement weight controlling the contribution of the boundary prior, and $\odot$ denotes element-wise multiplication.

$$B_l = \phi_{dilate}(E_p, r_l) - \phi_{erode}(E_p, r_l) \quad (21)$$

$$S_p^{(l)} = S_p^{(l-1)} \odot (1 - \lambda_b B_l) + \lambda_b B_l \quad (22)$$

2. Texture-Aware Interior Branch (TIB)

While BPB ensures boundary integrity, the Texture-Aware Interior Branch focuses on the lesion's internal consistency [39, 40]. Leveraging the gradient-enhanced feature maps F_g, an Adaptive Texture Smoothing (ATS) module removes irrelevant noise while preserving diagnostically significant patterns. The texture representation is obtained through Eq. (23) as follows [39]. Here T_m denotes the intermediate texture map, $\psi_{conv}(\cdot)$ represents a convolution operation for texture extraction, F_g is the gradient-aware feature map obtained from GEONet, S_p is the coarse segmentation mask, and $\odot$ denotes element-wise multiplication.

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

The ATS process applies a content-adaptive bilateral filter, as shown in Eq. (24) as follows [40]. Here, where $T_s(i,j)$ denotes the smoothed texture value at pixel location (i,j), G_s is the spatial Gaussian kernel, G_r is the range Gaussian kernel controlling texture preservation, Ωdenotes the local neighborhood window, and (p,q) represents neighboring pixel coordinates within Ω. The filtered texture map T_s is then masked to retain lesion regions, as shown in Eq. (25) as follows [38]. Where where S_t denotes the texture-refined segmentation mask and $\sigma(\cdot)$ is the sigmoid activation function that maps values to the range $[0,1]$.

$$T_m = \psi_{conv}(F_g \odot S_p) \quad (23)$$

$$T_s(i,j) = \frac{\sum_{(p,q)\in\Omega} G_s(\|(i,j)-(p,q)\|)\cdot G_r(|T_m(i,j)-T_m(p,q)|)\cdot T_m(p,q)}{\sum_{(p,q)\in\Omega} G_s(\|(i,j)-(p,q)\|)\cdot G_r(|T_m(i,j)-T_m(p,q)|)} \quad (24)$$

$$S_t = \sigma(T_s) \odot S_p \quad (25)$$

3. Synergistic Fusion Module (SFM)

The SFM merges the boundary-refined mask S_b and texture-refined mask S_t into the final segmentation output S_f. To balance boundary sharpness and texture uniformity, a Structure–Texture Consistency Loss is employed during training [28]. In Eq. (26), as follows [36], denotes the structure–texture consistency loss, S_{gt} is the ground-truth segmentation mask, $BCE(\cdot)$ represents the binary cross-entropy loss, $\nabla(\cdot)$ computes the spatial gradient operator, $\|\cdot\|_1$ denotes the L1 norm, and α, β, and γ are weighting coefficients controlling the contributions of the boundary, texture, and gradient consistency terms, respectively. The final mask is obtained via Eq. (27) as follows [32], where η is the learned fusion weight $(0 \le \eta \le 1)$, and S_f denotes the final refined segmentation mask. In Eq. (28) as follows [32], where S_p denotes the coarse segmentation mask generated by GEONet, E_p represents the predicted boundary (edge) map, F_g denotes the gradient-aware feature representation extracted by GEONet, $R_{BPB}(\cdot)$ is the Boundary Preservation Branch refinement operator, $R_{TIB}(\cdot)$ is the Texture-Aware Interior Branch refinement operator, and S_f is the final refined segmentation mask. Combining the above, the complete BESNet refinement process can be expressed as S_f.

$$L_{STC} = \alpha \cdot BCE(S_b, S_{gt}) + \beta \cdot BCE(S_t, S_{gt}) + \gamma \cdot \| \nabla S_b - \nabla S_t \|_1 \quad (26)$$

$$S_f = \eta \cdot S_b + (1-\eta) \cdot S_t \quad (27)$$

$$S_f = \eta \cdot R_{BPB}(S_p, E_p) + (1-\eta) \cdot R_{TIB}(S_p, F_g) \quad (28)$$

E. Implementation Details of GEONet and BESNet

The proposed GEONet–BESNet framework follows a lightweight encoder–refinement architecture optimized for retinal lesion segmentation. Input retinal images are resized to 512 × 512 pixels prior to processing. GEONet employs a five-stage convolutional encoder consisting of successive 3×3 convolutional layers with batch normalization and ReLU activation, where the number of feature channels progressively increases from 64 to 128, 256, 512, and finally 1024 filters to capture both local retinal textures and high-level lesion semantics. The GSA module is inserted after each encoder stage to enhance gradient-sensitive feature representations through spatial and channel attention mechanisms. The ECAL component incorporates a lightweight discriminator composed of convolution–batch normalization–LeakyReLU blocks operating on lesion boundary maps to enforce boundary realism. BESNet refines the coarse segmentation output using the Boundary Preservation Branch (BPB), Texture-Aware Interior Branch (TIB), and Synergistic Fusion Module (SFM). All convolutional operations utilize 3×3 kernels with same padding, while sigmoid activation generates pixel-level probability maps. The model is trained using the Adam optimizer with a learning rate of 1×10^{-4}, a batch size of 32, and early stopping with a patience of 15 epochs. The total loss combines segmentation, boundary, and adversarial objectives, with λedge = 0.30 and λadv = 0.20 determined through validation experiments.

IV. Results

Experiments were conducted using Python 3.10 with TensorFlow and PyTorch implementations. PyTorch Lightning was utilized for training management, while OpenCV, scikit-image, and NumPy supported preprocessing and data handling. All models were trained on a workstation equipped with an NVIDIA RTX A6000 GPU (48 GB). A batch size of 32, Adam optimizer, and an initial learning rate of 1×10^{-4} were employed. Early stopping with a patience of 15 epochs and dropout regularization were applied to reduce overfitting. Data augmentation included rotation, flipping, and intensity normalization. Performance evaluation was performed using stratified 5-fold cross-validation to ensure balanced and reliable assessment across datasets. To evaluate the robustness and generalizability of the proposed GEONet–BESNet framework, experiments were conducted on the DRD, APTOS 2019, and resized EyePACS datasets, which contain retinal fundus images acquired under diverse imaging conditions and varying DR severities. Prior to training, all images underwent illumination normalization, CLAHE-based contrast enhancement, noise suppression, and region-of-interest extraction. Images exhibiting severe unreadable retinal structures were filtered during preprocessing. The processed datasets were partitioned into training (70%), validation (15%), and testing (15%) subsets while preserving dataset-specific class distributions.

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

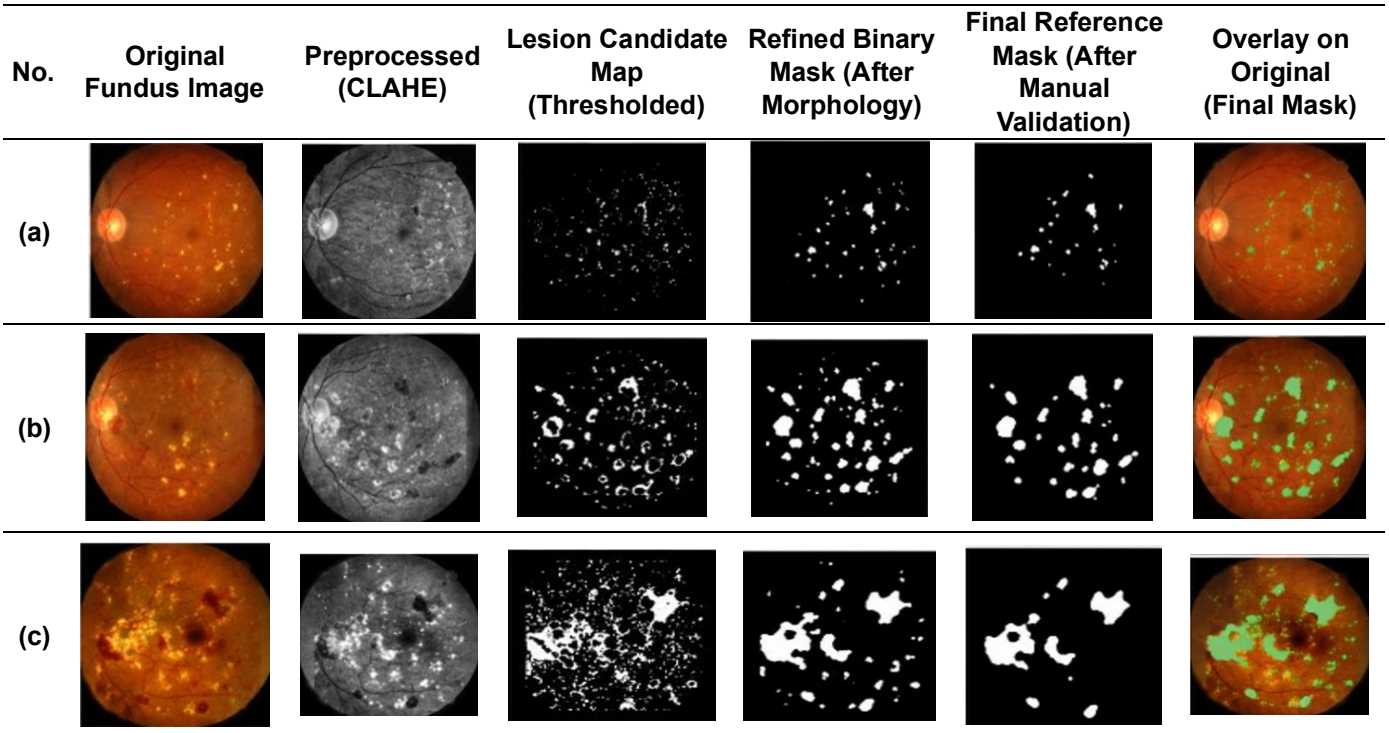

Fig. 2. Representative Examples of Semi-Automatic Lesion Mask Generation and Validation (a) Example 1 (Mild DR), (b) Example 2 (Moderate DR), (c) Example 3 (Severe DR)

Table 2. Analysis of Region Overlap and Boundary Fidelity.

Datasets	Methods	Dice Coefficient	IoU	Boundary Accuracy	Hausdorff Distance
Dataset 1 (DRD)	WE-DNN	0.90	0.86	0.88	5.3
	VGGNet	0.85	0.83	0.84	6.6
	GAMDL	0.87	0.85	0.85	6.0
	CNbDL	0.83	0.79	0.82	7.8
	EDLM	0.81	0.78	0.79	9.1
	Proposed Work	0.95	0.89	0.91	3.8
Dataset 2 (APTOS)	WE-DNN	0.83	0.82	0.83	5.6
	VGGNet	0.79	0.79	0.77	7.0
	GAMDL	0.81	0.80	0.79	6.1
	CNbDL	0.78	0.77	0.75	8.2
	EDLM	0.77	0.76	0.73	9.9
	Proposed Work	0.87	0.85	0.86	4.3
Dataset 3 (EyePACS)	WE-DNN	0.86	0.84	0.87	5.7
	VGGNet	0.82	0.80	0.79	7.1
	GAMDL	0.84	0.82	0.83	6.3
	CNbDL	0.82	0.78	0.79	7.9
	EDLM	0.79	0.77	0.77	8.7
	Proposed Work	0.89	0.87	0.89	5.1

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

Data augmentation was applied exclusively to the training set to mitigate class imbalance and improve model generalization. It should be noted that the original DRD, APTOS 2019, and EyePACS datasets primarily provide image-level diabetic retinopathy grading labels rather than native pixel-level lesion annotations. Therefore, lesion masks used for supervised segmentation training and quantitative evaluation were obtained through an additional annotation generation and validation procedure. These reference lesion masks served as the ground truth for the computation of Dice Coefficient, IoU, Boundary Accuracy, Hausdorff Distance, DSC, and SSIM.

Table 3. Analysis of Structural and Patch-Level Consistency.

Datasets	Methods	Dice Similarity Coefficient (DSC) %	Structural Similarity Index Measure (SSIM) %	Proportion of correct patches %
Dataset 1 (DRD)	WE-DNN	87.7	88.5	91.4
	VGGNet	85.5	85.7	87.5
	GAMDL	86.7	87.5	89.4
	CNbDL	82.1	82.5	81.8
	EDLM	79.9	81.6	80.9
	Proposed Work	89.9	91.9	93.9
Dataset 2 (APTOS)	WE-DNN	89.5	88.8	91.9
	VGGNet	84.8	85.3	87.4
	GAMDL	87.9	88.5	90.6
	CNbDL	83.8	84.7	85.8
	EDLM	79.7	82.9	80.6
	Proposed Work	91.6	90.5	95.4
Dataset 3 (EyePACS)	WE-DNN	88.7	89.2	89.3
	VGGNet	86.3	86.4	88.8
	GAMDL	87.9	89.3	92.7
	CNbDL	82.3	83.3	82.4
	EDLM	80.4	83.9	82.3
	Proposed Work	90.5	90.3	94.5

A. Lesion Mask Generation and Validation

Although the DRD, APTOS 2019, and EyePACS datasets primarily provide image-level diabetic retinopathy severity grades rather than pixel-wise lesion annotations, the proposed segmentation framework requires lesion-level masks for supervised training and quantitative evaluation. Therefore, lesion masks were generated through a semi-automatic annotation pipeline. Initially, retinal lesions were localized using contrast-enhanced candidate extraction based on CLAHE preprocessing, morphological top-hat filtering, and adaptive thresholding. Candidate lesion regions were subsequently refined using connected-component analysis and boundary smoothing operations. To improve annotation quality, all generated masks underwent manual verification and correction by two independent reviewers with experience in retinal image analysis. Disagreements were resolved through consensus review. The resulting lesion masks were utilized exclusively for segmentation training and evaluation. Representative examples of generated masks are provided in Fig. 2. Since these annotations were generated specifically for this study, they are not part of the original public datasets.

B. Fairness and Reproducibility of Comparative Evaluation

The comparative evaluation was conducted under a unified experimental protocol to ensure fairness and reproducibility across all methods. The baseline models, including WE-DNN, GAMDL, VGGNet, CNbDL, and EDLM, were implemented and evaluated on the same datasets, with the same preprocessing procedures, augmentation strategies, and data partitioning scheme as the proposed GEONet–BESNet framework. Specifically, all retinal images underwent identical illumination normalization, CLAHE-based contrast enhancement, noise suppression, and region-of-interest extraction prior to training. Furthermore, the same training, validation, and testing subsets were

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

maintained for all methods to eliminate performance variations arising from dataset partition differences. To ensure consistency, all models were trained within the same computational environment and evaluated using identical performance metrics. Hyperparameter settings for the baseline methods were initialized according to the configurations reported in their original studies and subsequently optimized using the same validation protocol employed for the proposed framework. Data augmentation operations, including rotation, horizontal flipping, and intensity normalization, were also applied uniformly across all models. Consequently, the reported performance differences reflect the intrinsic capabilities of the respective architectures rather than variations in preprocessing, training conditions, or experimental setup. This standardized evaluation protocol improves the transparency, fairness, and reproducibility of the comparative analysis.

Table 4. Pixel-level sensitivity, specificity, and precision comparison for lesion segmentation on the DRD, APTOS, and EyePACS datasets

Authors	Datasets	Methods	Sensitivity (Recall)	Specificity	Precision (Positive Predictive Value)
Nazir et al. [17]	Dataset 1 (DRD)	WE-DNN	89.9	88.9	91.9
Jabbar et al. [19]		VGGNet	84.3	85.2	86.4
Zedadra et al. [18]		GAMDL	88.4	88.7	90.2
Govindharaj et al. [20]		CNbDL	81.9	82.5	81.9
Banupriya et al. [21]		EDLM	79.3	80.8	78.4
Proposed Work		GEONet + BESNet	91.9	90.4	95.8
Nazir et al. [17]	Dataset 2 (APTOS)	WE-DNN	90.3	87.7	92.7
Jabbar et al. [19]		VGGNet	86.5	84.3	86.9
Zedadra et al. [18]		GAMDL	88.6	86.5	89.2
Govindharaj et al. [20]		CNbDL	86.2	82.5	85.6
Banupriya et al. [21]		EDLM	85.4	80.8	82.9
Proposed Work		GEONet + BESNet	94.9	93.7	97.3
Nazir et al. [17]	Dataset 3 (EyePACS)	WE-DNN	88.6	87.5	85.7
Jabbar et al. [19]		VGGNet	85.5	82.4	84.3
Zedadra et al. [18]		GAMDL	87.7	83.3	85.6
Govindharaj et al. [20]		CNbDL	84.5	82.5	80.3
Banupriya et al. [21]		EDLM	78.9	80.7	78.9
Proposed Work		GEONet + BESNet	92.5	94.6	92.8

C. Region Overlap and Boundary Fidelity Analysis

Segmentation performance was evaluated using Dice Coefficient, IoU, Boundary Accuracy, and Hausdorff Distance, which collectively assess region overlap and boundary fidelity (Table 2). The proposed GEONet–BESNet framework consistently achieved superior results across all datasets. On DRD, it attained Dice, IoU, and Boundary Accuracy values of 0.95, 0.89, and 0.91, respectively, outperforming all competing methods. Similar improvements were observed on APTOS (0.87, 0.85, 0.86) and EyePACS (0.89, 0.87, 0.89). In terms of boundary preservation, the framework achieved the lowest Hausdorff Distances of 3.8, 4.3, and 5.1 on DRD, APTOS, and EyePACS, respectively. These findings demonstrate that integrating GEONet and BESNet enhances lesion localization, structural consistency, and boundary delineation across heterogeneous retinal imaging datasets.

D. Region Overlap and Boundary Fidelity Analysis

To further evaluate segmentation quality, DSC, SSIM, and PCP metrics were assessed as shown in Table 4. The proposed GEONet–BESNet framework consistently achieved superior structural and local consistency across all datasets. On DRD, it attained DSC, SSIM, and PCP values of 89.9%, 91.9%, and 93.9%, respectively. Corresponding values on APTOS were 91.6%, 90.5%, and 95.4%, while EyePACS achieved 90.5%, 90.3%, and 94.5%, as shown in Table 3. These results exceeded those of competing methods, demonstrating improved lesion delineation, structural fidelity, and patch-level robustness. The consistent gains across heterogeneous datasets indicate that integrating gradient-aware segmentation with boundary-enhanced refinement effectively

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

preserves lesion structures and maintains reliable segmentation performance under varying retinal imaging conditions.

V. Discussions

A. Evaluation of Statistical and Clinical Relevance

Table 4 demonstrates that the proposed GEONet–BESNet framework consistently outperformed the compared studies across the DRD, APTOS, and EyePACS datasets in terms of Sensitivity, Specificity, and Precision. Compared with the WE-DNN framework of Nazir et al. [17], the proposed method achieved higher sensitivity on all three datasets, indicating improved detection of subtle lesion pixels while reducing false negatives. Similarly, compared with the VGGNet-based approach of Jabbar et al. [19], the proposed framework showed substantial improvements in specificity and precision, suggesting more reliable discrimination between lesion and non-lesion regions with fewer false-positive predictions. Although the GAMDL model proposed by Zedadra et al. [18] demonstrated competitive performance, GEONet–BESNet consistently achieved higher diagnostic metrics across all datasets, reflecting the effectiveness of combining gradient-guided feature extraction with boundary-aware refinement. The CNbDL framework of Govindharaj et al. [20] and the EDLM model of Banupriya et al. [21] exhibited comparatively lower performance, particularly on the heterogeneous EyePACS dataset, indicating limited robustness to imaging variability. In contrast, the proposed framework maintained consistently high performance across all three datasets, demonstrating superior generalization capability. These findings are consistent with previous studies emphasizing the importance of boundary preservation and feature enhancement for retinal lesion segmentation [18], [20],

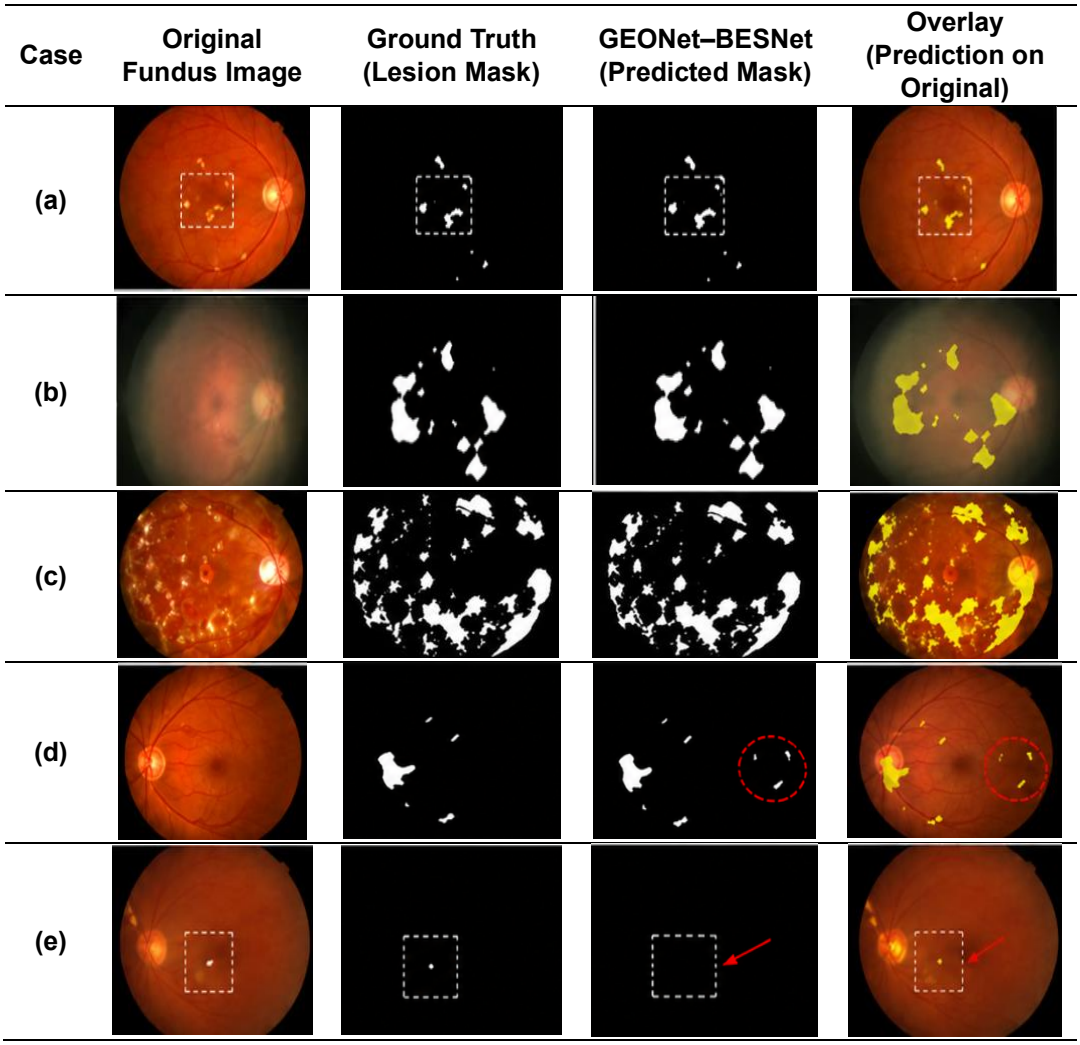

Fig. 3. Qualitative segmentation results (a) Small Lesion Detection, (b) Blurred Image, (c) Severe DR, (d) False Positive Example, (e) False Negative Example

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

and further demonstrate that integrating GEONet and BESNet improves structural consistency, lesion localization, and diagnostic reliability across diverse retinal imaging conditions.

B. Comparative Insights

Comparative evaluation across the DRD, APTOS, and EyePACS datasets demonstrates the superiority of the proposed GEONet–BESNet framework over WE-DNN, VGGNet, GAMDL, CNbDL, and EDLM. The framework consistently achieved higher DSC, SSIM, Sensitivity, Specificity, and Precision values, indicating improved structural fidelity, lesion delineation, and segmentation reliability. Unlike competing methods that often exhibit trade-offs between overlap accuracy, boundary preservation, and diagnostic performance, the proposed approach maintains balanced performance across all evaluation metrics. Its consistent results across heterogeneous datasets further highlight its robustness and generalization capability, confirming the effectiveness of integrating gradient-aware segmentation with boundary-enhanced refinement for accurate retinal lesion segmentation. To complement the quantitative evaluation, qualitative segmentation examples under multiple retinal imaging conditions are presented in Fig. 3. The proposed framework accurately delineates small lesion regions and maintains boundary fidelity even in low-contrast and blurred images. In severe diabetic retinopathy cases, the model successfully captures extensive pathological regions while preserving lesion morphology. Nevertheless, several challenging examples reveal occasional false-positive activations near vessel structures and false-negative omissions in extremely low-contrast lesions. These observations indicate that although the proposed framework demonstrates strong robustness across heterogeneous retinal images, segmentation performance remains influenced by lesion visibility and image quality.

C. Implications of the Study

The proposed GEONet–BESNet framework has important implications for both clinical practice and automated ophthalmic image analysis. By improving lesion boundary preservation and structural consistency, the framework can support ophthalmologists in more reliable identification of microaneurysms, hemorrhages, and exudates, particularly during early-stage diabetic retinopathy screening. Its consistent performance across the DRD, APTOS, and EyePACS datasets indicates strong potential for deployment in heterogeneous clinical environments where image quality and acquisition conditions vary considerably. Furthermore, the lightweight dual-stage architecture provides a practical alternative to computationally intensive ensemble and multimodal methods, making it more suitable for resource-constrained healthcare settings and large-scale screening programs. The proposed gradient-aware and boundary-enhanced segmentation strategy may also be extended to other retinal diseases requiring accurate lesion delineation, including glaucoma, age-related macular degeneration, and retinal vein occlusion. Overall, the study contributes a clinically interpretable and scalable segmentation framework that can facilitate future computer-aided diagnosis systems and support more efficient retinal disease screening.

D. Ablation Analysis of GEONet and BESNet

To quantify the contribution of each proposed component, an ablation study was conducted by progressively enabling the modules of the proposed framework. The baseline model consisted of the segmentation backbone without GSA, ECAL, or BESNet. Subsequently, GSA and ECAL were introduced independently to evaluate their individual effects on feature discrimination and boundary preservation, followed by the complete GEONet architecture and BESNet refinement stage. The results demonstrate a consistent improvement across all segmentation metrics as additional components are incorporated. The baseline backbone achieved a Dice score of 84.6%, which improved to 86.8% after introducing GSA and further increased to 88.1% with ECAL. The complete GEONet architecture achieved a Dice score of 89.2%, highlighting the benefits of attention-guided feature extraction and edge-aware optimization. When BESNet was applied as a standalone refinement stage, structural consistency further improved, yielding a Dice score of 89.8% and better boundary preservation. Finally, the integrated GEONet–BESNet framework achieved the highest overall performance with average Dice, IoU, and Boundary Accuracy values of 90.3%, 87.0%, and 88.7%, respectively, confirming the complementary contributions of gradient-guided attention, adversarial boundary supervision, and boundary–texture refinement.

E. Limitations and Future Scope

Although the proposed framework demonstrated strong segmentation performance, a detailed computational complexity analysis, including parameter count, FLOPs, GPU memory consumption, training time, and inference latency, was not conducted in the current study. Consequently, deployment feasibility should be interpreted based on architectural design considerations rather than quantitative efficiency measurements. Comprehensive benchmarking against baseline models under identical hardware settings will be investigated in future work. Future work will focus on improving robustness, domain adaptation, minority-lesion representation,

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

explainable segmentation, and validation on larger multi-center datasets.

V. Conclusion

This study presented the GEONet–BESNet framework for retinal lesion segmentation in diabetic retinopathy images, integrating gradient-guided feature extraction with boundary-aware refinement. The proposed approach consistently outperformed the compared methods across the DRD, APTOS, and EyePACS datasets, achieving average Dice, IoU, and Boundary Accuracy values of 90.3%, 87.0%, and 88.7%, respectively, while maintaining strong structural consistency and accurate lesion delineation. The results demonstrate that the combination of gradient-aware segmentation and boundary refinement enhances lesion boundary preservation and robustness under diverse retinal imaging conditions. Although improved lesion segmentation may support downstream clinical assessment workflows, the present study focuses exclusively on pixel-level lesion segmentation and does not evaluate image-level diabetic retinopathy grading or patient-level diagnosis.

Acknowledgment

The authors gratefully acknowledge financial support from Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology under the Research Development Fund (RDF), Grant No. VTU/RDF/FY2026-27/R03. This support was instrumental in facilitating the successful execution of this research work.

Funding

This research was financially supported by the Research Development Fund (RDF), Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, under Grant No. VTU/RDF/FY2026-27/R03.

Data Availability

The retinal image datasets used in this study are publicly available through the Kaggle platform. The Diabetic Retinopathy Detection (DRD) dataset is available at https://www.kaggle.com/c/diabetic-retinopathy-detection, the APTOS 2019 Blindness Detection dataset is available at https://www.kaggle.com/c/aptos2019-blindness-detection, and the resized EyePACS Diabetic Retinopathy dataset is available at https://www.kaggle.com/datasets/mohlamin/resized-eyepacs-diabetic-retinopathy-dataset. The lesion masks used for supervised segmentation were generated by the authors through a semi-automatic annotation and manual validation process. The generated lesion masks, annotation protocol, and preprocessing details are available from the corresponding author upon reasonable request.

Author Contribution

P. Malathi contributed to conceptualization and methodology. D. Dhinakaran contributed to original draft preparation, data curation, and supervision. S. Jeyalakshmi contributed to methodology and validation. P. Kavitha contributed to visualization and investigation. S. Palpandi contributed to methodology, review, and editing. G. Prabaharan contributed to software development, model implementation, and validation. All authors reviewed and approved the final manuscript.

Declarations

Ethical Approval

This study used publicly available anonymized retinal image datasets and did not involve direct interaction with human participants or animals. Therefore, formal ethical approval was not required. All datasets were used in accordance with their respective access conditions and terms of use.

Consent for Publication Participants.

Not applicable. The study used publicly available anonymized datasets and did not involve the collection of identifiable participant information.

Competing Interests

The authors declare no competing interests.

References

[1] R. Dahiya, N. Agarwal, S. Singh, D. Verma, and S. Gupta, “Diabetic Retinopathy Eye Disease Detection Using Machine Learning”, *EAI Endorsed Trans IoT*, vol. 10, Mar. 2024. doi:10.4108/eetiot.5349

[2] M. Asif, F. Ur Rehman, Z. Rashid, A. Hussain, A. Mirza and W. S. Qureshi, "An Insight on the Timely Diagnosis of Diabetic Retinopathy Using Traditional and AI-Driven Approaches," *in IEEE Access*, vol. 13, pp. 116869-116886, 2025, doi: 10.1109/ACCESS.2025.3583647.

[3] A. A. Khan, K. M. Ahmad, S. Shafiq, M. U. Akram, and J. Shao, “ATLASS: An anatomically-aware self-supervised learning framework for generalizable retinal disease detection,” *IEEE Journal of Biomedical and Health Informatics*, 2025.

[4] B. Hassan, H. Raja, T. Hassan, M. U. Akram, H. Raja, A. A. Abdalrazaq, S. Yousefi, and N. Werghi, “A comprehensive review of artificial

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

intelligence models for screening major retinal diseases," *Artificial Intelligence Review*, vol. 57, no. 5, p. 111, 2024.

[5] A. Tulsani, P. Kumar, and S. Pathan, "Automated segmentation of optic disc and optic cup for glaucoma assessment using improved UNETþþ architecture," *Biocybern. Biomed. Eng*., vol. 41, no. 2, pp. 819–832, 2021, doi: 10.1016/j.bbe.2021.05.011.

[6] W. T. Song, I. C. Lai, and Y. Z. Su, "A statistical robust glaucoma detection framework combining Retinex, CNN, and DOE using fundus images," *IEEE Access*, vol. 9, pp. 103772–103783, 2021, doi: 10.1109/ACCESS.2021.3098032.

[7] G. A. Saleh, N. M. Batouty, S. Haggag, A. Elnakib, F. Khalifa, F. Taher, M. A. Mohamed, R. Farag, H. Sandhu, A. Sewelam, and A. El-Baz, "The role of medical image modalities and ai in the early detection, diagnosis and grading of retinal diseases: a survey," *Bioengineering*, vol. 9, no. 8, p. 366, 2022.

[8] S. A. Khowaja, K. Dev, S. M. Anwar, and M. G. Linguraru, "Selffed: Self-supervised federated learning for data heterogeneity and label scarcity in medical images," *Expert Syst. Appl*., vol. 261, p. 125493, 2025

[9] S. Parsa and T. Khatibi, "Grading the severity of diabetic retinopathy using an ensemble of self-supervised pre-trained convolutional neural networks: Essp-cnns," *Multimedia Tools and Applications*, vol. 83, pp. 89 837–89 870, 2024.

[10] O. Kovalyk, J. Morales-Sanchez, R. Verd ´ u-Monedero, I. Sell ´ es-Navarro, ´ A. Palazon-Cabanes, and J.-L. Sancho-G ´ omez, "Papila: Dataset with ´ fundus images and clinical data of both eyes of the same patient for glaucoma assessment," *Scientific Data*, vol. 9, no. 1, p. 291, 2022.

[11] N. Mukherjee, S. Sengupta, M. Nadeem Ahmed, S. Irfan Yaqoob, M. Rashid Hussain and A. Taha Zamani, "Bi-Directional Hybrid Attention Feature Pyramid Network for Detecting Diabetic Macular Edema in Retinal Fundus Images," *in IEEE Access*, vol. 13, pp. 38726-38744, 2025, doi: 10.1109/ACCESS.2025.3545873.

[12] D. Yi, P. Baltov, Y. Hua, S. Philip, and P. K. Sharma, "Compound scaling encoder-decoder (CoSED) network for diabetic retinopathy related bio-marker detection," *IEEE Journal of Biomedical and Health Informatics*, early access, Sep. 11, 2023, doi: 10.1109/JBHI.2023.3313785.

[13] L. Dai, L. Wu, H. Li, C. Cai, Q. Wu, H. Kong, R. Liu, X. Wang, X. Hou, Y. Liu, X. Long, Y. Wen, L. Lu, Y. Shen, Y. Chen, D. Shen, X. Yang, H. Zou, B. Sheng, *et al*., "A deep learning system for detecting diabetic retinopathy across the disease spectrum," *Nature Communications*, vol. 12, no. 1, pp. 1–11, 2021, doi: 10.1038/s41467-021-23458-5.

[14] A. Asia, C. Zhu, S. A. Althubiti, Y. Xiao, P. Ouyang, and M. A. A., "Detection of diabetic retinopathy in retinal fundus images using CNN classification models," *Electronics*, vol. 11, no. 17, p. 2740, 2022, doi: 10.3390/electronics11172740.

[15] A. Malhi, R. Grewal, and H. S. Pannu, "Detection and diabetic retinopathy grading using digital retinal images," *International Journal of Intelligent Robotics and Applications*, vol. 7, pp. 426–458, 2023, doi: 10.1007/s41315-022-00269-5.

[16] P. Modi and Y. Kumar, "Smart detection and diagnosis of diabetic retinopathy using bat based feature selection algorithm and deep forest technique," *Computers & Industrial Engineering*, vol. 182, p. 109364, 2023, doi: 10.1016/j.cie.2023.109364.

[17] K. Nazir, J. Kim and Y. -C. Byun, "Enhancing Early-Stage Diabetic Retinopathy Detection Using a Weighted Ensemble of Deep Neural Networks," *in IEEE Access*, vol. 12, pp. 113565-113579, 2024, doi: 10.1109/ACCESS.2024.3432867.

[18] A. Zedadra, O. Zedadra, M. Yassine Salah-Salah and A. Guerrieri, "Graph-Aware Multimodal Deep Learning for Classification of Diabetic Retinopathy Images," *in IEEE Access,* vol. 13, pp. 74799-74810, 2025, doi: 10.1109/ACCESS.2025.3564529.

[19] A. Jabbar, S. Naseem, J. Li, *et al*., "Deep transfer learning-based automated diabetic retinopathy detection using retinal fundus images in remote areas," *International Journal of Computational Intelligence Systems*, vol. 17, p. 135, 2024, doi: 10.1007/s44196-024-00520-w.

[20] I. Govindharaj, R. Rampriya, G. Michael, *et al*., "Capsule network-based deep learning for early and accurate diabetic retinopathy detection," *International Ophthalmology*, vol. 45, p. 78, 2025, doi: 10.1007/s10792-024-03391-4.

[21] Banupriya, V., Anusuya, S., "Strategy for Rapid Diabetic Retinopathy Exposure Based on Enhanced Feature Extraction Processing," *Computers, Materials & Continua*, 75(3), 5597–5613, 2023. doi:10.32604/cmc.2023.038696

[22] Gour, N., & Khanna, P., "Automated glaucoma detection using GIST and pyramid histogram of

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

oriented gradients (PHOG) descriptors," *Pattern Recognition Letters*, 137, 3-11, 2020. https://doi.org/10.1016/j.patrec.2019.04.004.

[23] T. M. Devi, P. Karthikeyan, B. M. Kumar, and M. Manikandakumar, "Diabetic retinopathy detection via deep learning based dual features integrated classification model," *Technology and Health Care*, 2025, doi: 10.1177/09287329241292939.

[24] S. Rajeshwar, S. Thaplyal, A. M., and S. S. G., "Diabetic retinopathy detection using DL-based feature extraction and a hybrid attention-based stacking ensemble," *Advances in Public Health*, vol. 2025, no. 1, p. 8863096, 2024, doi: 10.1155/adph/8863096.

[25] R. Nandakumar, P. Saranya, V. Ponnusamy, S. Hazra, and A. Gupta, "Detection of diabetic retinopathy from retinal images using DenseNet models," *Computer Systems Science and Engineering*, vol. 45, no. 1, pp. 279–292, 2023, doi: 10.32604/csse.2023.028703.

[26] G. Sun, X. Wang, L. Xu, C. Li, W. Wang, Z. Yi, H. Luo, Y. Su, J. Zheng, Z. Li, Z. Chen, H. Zheng, and C. Chen, "Deep learning for the detection of multiple fundus diseases using ultra-widefield images," *Ophthalmology and therapy*, vol. 12, no. 2, pp. 895–907, 2023.

[27] J. Y. Choi, I. H. Ryu, J. K. Kim, I. S. Lee, and T. K. Yoo, "Development of a generative deep learning model to improve epiretinal membrane detection in fundus photography," *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, p. 25, 2024

[28] F. Jeribi, T. Nazir, M. Nawaz, A. Javed, M. Alhameed, and A. Tahir, "Recognition of diabetic retinopathy and macular edema using deep learning," *Medical & Biological Engineering & Computing*, pp. 1–15, 2024.

[29] B. J. Bhatkalkar, D. R. Reddy, S. Prabhu, and S. V. Bhandary, "Improving the performance of convolutional neural network for the segmentation of optic disc in fundus images using attention gates and conditional random fields," IEEE Access, vol. 8, pp. 29299–29310, 2020, doi: 10.1109/ACCESS.2020.2972318.

[30] D. S. S. Raja, S. Kumarganesh, K. M. Sagayam, and H. Dang, "Diabetic Retinopathy Detection and Grading System Using Deep Learning Approach," *SAGE Open Medicine*, 2026, doi: 10.1177/20552076251410982.

[31] M. Alam, E. J. Zhao, C. K. Lam, and D. L. Rubin, "Segmentation-Assisted Fully Convolutional Neural Network Enhances Deep Learning Performance to Identify Proliferative Diabetic Retinopathy," *Journal of Clinical Medicine*, vol. 12, no. 1, 2023, doi: 10.3390/jcm12010385.

[32] S. Das, K. Kharbanda, M. Suchetha, R. Raman, and E. Dhas, "Deep Learning Architecture Based on Segmented Fundus Image Features for Classification of Diabetic Retinopathy," *Biomedical Signal Processing and Control*, vol. 68, 2021, doi: 10.1016/j.bspc.2021.102600.

[33] A. Sebastian, O. Elharrouss, S. Al-Maadeed, and N. Almaadeed, "A Survey on Diabetic Retinopathy Lesion Detection and Segmentation," *Applied Sciences*, vol. 13, no. 8, 2023, doi: 10.3390/app13085111.

[34] R. Raman, S. Srinivasan, S. Virmani, et al., "Fundus Photograph-Based Deep Learning Algorithms in Detecting Diabetic Retinopathy," *Eye*, vol. 33, 2019, doi: 10.1038/s41433-018-0269-y.

[35] A. Porwal et al., "IDRiD: Diabetic Retinopathy – Segmentation and Grading Challenge," *Medical Image Analysis*, vol. 59, 2020, doi: 10.1016/j.media.2019.101561.

[36] S. Dinesen, M. G. Schou, C. V. Hedegaard, et al., "A Deep Learning Segmentation Model for Detection of Active Proliferative Diabetic Retinopathy," *Ophthalmology and Therapy*, vol. 14, pp. 1053–1063, 2025, doi: 10.1007/s40123-025-01127-w.

[37] N. Sharma and P. Lalwani, "A Multi Model Deep Net with an Explainable AI Based Framework for Diabetic Retinopathy Segmentation and Classification," *Scientific Reports*, vol. 15, 2025, doi: 10.1038/s41598-025-93376-9.

[38] A. Rehman, G. Naijie, S. Ojo, et al., "FISM: Harnessing Deep Learning and Reinforcement Learning for Precision Detection of Microaneurysms and Retinal Exudates for Early Diabetic Retinopathy Diagnosis," *BioData Mining*, vol. 18, 2025, doi: 10.1186/s13040-025-00485-2.

[39] R. Nadda, J. Singh, and U. Shrivastava, "Automatic Diabetic Retinopathy Detection Using an Ensemble Learning Approach and Classifiers with Self-Adjusting Weights," *Soft Computing*, vol. 29, 2025, doi: 10.1007/s00500-025-10773-y.

[40] I. Govindharaj, A. Poongodai, G. Rajaram, et al., "Enhanced Diabetic Retinopathy Detection Using U-Shaped Network and Capsule Network-Driven Deep Learning," *MethodsX*, vol. 14, 2025, doi: 10.1016/j.mex.2024.103052.

Author Biography

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734

Ms. P. Malathi is working as an assistant professor in the Department of CSE at the SRM Institute of Science and Technology, Ramapuram, Chennai. She completed her master's degree in the specialization of Computer Science and Engineering from Anna University, Chennai, in 2013 and is currently pursuing her PhD from Annamalai University, Chidambaram. She has 10 years of teaching experience and expertise in programming languages such as Python, Java, and C++, and web technologies like HTML, CSS, and JavaScript. She has published articles in national and international journals and conferences. She has been a professional member of reputed technical societies.

Dr. D. Dhinakaran is an Associate Professor in the Department of Computer Science and Engineering at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai. He holds a B.E., M.E., and Ph.D. in Computer Science from Anna University. His research expertise includes Privacy-Preserving Data Mining, Artificial Intelligence, Internet of Things (IoT), MANETs, Cloud Computing, and Image Processing. A prolific researcher, he has presented over 30 papers at international conferences, published more than 60 articles in reputed international journals, and filed three patents. He has been honored as one of the Top 2% Scientists globally by Elsevier and Stanford University. He actively supervises Ph.D. scholars and serves as a reviewer for reputable international journals and conferences, reflecting his strong commitment to academic excellence and research advancement. His innovative contributions and dedication to academia have made a significant impact on the evolving fields of Computer Science and Engineering.

Mrs. S. Jeyalakshmi is working as an Assistant Professor, Department of Artificial Intelligence and Data Science in P.S.R. Engineering College, Sivakasi. She completed her M.E. in Computer Science and Engineering from P.S.R Engineering College, Sivakasi, in the year 2011. She obtained her B.E in Computer Science and Engineering in the year 2009 from P.S.R Engineering College, Sivakasi. She has 3 years of teaching experience. Her research interests are Data Science, Machine Learning, and Deep Learning.

Dr. P. Kavitha, Associate Professor, Department of Computer Science and Engineering, R.M.K. Engineering College, Kavarapettai, Tiruvallur. She received her B.E. Degree from the University of Madras in 1999 and her M.E. Degree from Anna University, Chennai in 2007. She received her Doctoral Degree from Anna University, Chennai in 2020. Her research interests include Image processing, Artificial Intelligence, and Machine Learning. She is a life member of ISTE. She received her Academic Excellence Award by the Institution of Engineers in March 2023. Professional Membership: IEEE – Annual Member (2026)

Dr. S. Palpandi serves as Assistant Professor-Senior Grade in the Department of Computer Science and Engineering at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai. He holds a Ph.D. in Information and Communication Systems from Anna University, along with undergraduate and postgraduate degrees in Electronics and Communication Engineering and Computer Science and Engineering. With over 15 years of academic experience at reputable institutions, his research primarily focuses on Image Processing applications in healthcare, with emphasis on Electronic Health Records, Data Security, Health Data Analytics, and Networking Communications. He has published several articles in SCI(E)-indexed journals, international conferences, and UGC Care-listed journals, and he also holds two international patents.

Dr. G. Prabaharan is working as an Associate Professor at the Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai. He has completed an Ph.D. in Computer Science and Engineering, M.Tech. Degree in Computer Science and Engineering, B.E Degree in Computer Science and Engineering. His research areas of interest include machine learning, IoT, Cloud computing, Edge computing, and WSN. He has a total of 15 Years of Teaching and research experience in various Engineering Colleges. He is acting as a reviewer for journals; he has also published papers in various International and National Journals. He has published several papers in well-known peer-reviewed journals.

Manuscript received 6 May 2026; Revised 10 July 2026; Accepted 20 July 2026; Available online 10 August 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i4.1734