Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 9, No. 1, January 2027, pp: 1-14 **e-ISSN: 2656-8632**

RESEARCH ARTICLE **OPEN ACCESS**

Neuro-Spectral Causal Fusion Network with Lesion Dependency Reasoning for Explainable Diabetic Retinopathy Detection

K. Venu[1], B. Thanga Parvathi[2], Kunchanapalli Rama Krishna[3], and Yuvaraj S.[4]

[1] Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, India.
[2] Department of Computer Technology, Bannari Amman Institute of Technology, Tamil Nadu, India.
[3] Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Andhra Pradesh, India.
[4] Department of Computer Science and Engineering at Sri Eshwar College of Engineering, Coimbatore, India.

Corresponding author: K. Venu (e-mail: venu.kalaimagal@gmail.com), **Author(s) Email**: Thanga Parvathi (e-mail: drbtparvathi@gmail.com), Kunchanapalli Rama Krishna (e-mail: tenalirama@kluniversity.in), Yuvaraj S (e-mail: yuvaraj.scse@sece.ac.in)

Abstract

Background: Diabetic Retinopathy (DR) is a progressive microvascular complication of diabetes and a leading cause of preventable vision impairment worldwide. Early and accurate identification of DR severity is essential for timely intervention and effective disease management. However, conventional retinal screening relies heavily on manual examination of fundus images, which is labor-intensive, subjective, and dependent on specialist expertise. Although recent deep-learning models have demonstrated promising diagnostic performance, many existing approaches inadequately capture lesion-level relationships, underutilize clinical information, and provide limited interpretability for clinical decision-making.

Aims: this study proposes a novel Neuro-Spectral Causal Fusion Network (NSCF-Net) for explainable diabetic retinopathy detection and severity grading.

Methods: The proposed framework integrates an adaptive Neuro-Spectral Enhancement Module to improve lesion visibility, a Lesion Dependency Reasoning Graph to model pathological interactions among retinal abnormalities, and a Dual-Path Hybrid Transformer to jointly learn local lesion characteristics and global retinal context. Furthermore, a Causal Multimodal Fusion Unit combines retinal image representations with clinically relevant parameters, including HbA1c level, blood pressure, age, and duration of diabetes. To improve transparency and clinical trustworthiness, an Explainable Lesion Attribution Module generates counterfactual attention maps and hierarchical saliency visualizations highlighting diagnostically significant retinal regions. The proposed framework was evaluated on the APTOS 2019 diabetic retinopathy dataset, consisting of 3,662 retinal fundus images distributed across five DR severity classes.

Results: Experimental results demonstrate the effectiveness of NSCF-Net, achieving 99.1% accuracy, 98.9% precision, 99.0% recall, and 98.8% F1-score. Comparative analyses and ablation studies further confirm the contribution of neuro-spectral enhancement, lesion dependency reasoning, multimodal fusion, and explainability components to overall performance improvement.

Conclusion: These findings indicate that NSCF-Net provides an accurate, interpretable, and clinically reliable framework for automated diabetic retinopathy screening, severity assessment, and computer-aided ophthalmic decision support.

Keywords Diabetic Retinopathy Detection, Explainable Artificial Intelligence, Multimodal Deep Learning, Lesion Dependency Reasoning, Hybrid Transformer Network.

I. Introduction

Diabetic Retinopathy (DR) is one of the major complications of Diabetes mellitus and the primary cause of preventable blindness all over the world. It occurs as a result of chronic hyperglycaemia, which causes the blood vessels of the retina to become damaged and causes pathological changes in the retina, including microaneurysms, haemorrhages, exudates and neovascularization [1]. Early detection can prevent severe vision loss. Manual diagnosis of retinopathy from retinal fundus images is reliable but time-consuming, subjective and requires experience of ophthalmologists, so it is important to develop automated computer-aided diagnostic systems [2].

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Over the past few years, the paradigms of deep learning, such as Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and hybrid models, have made remarkable strides in the field of DR detection.In recent years, deep learning has improved significantly in DR detection, particularly with the development of Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and hybrid models. Yet, the majority of the current approaches are mainly concerned with image classification research and lack lesion dependencies, multimodal clinical data and model interpretability. In addition, many of the deep learning models are black boxes, which hinders clinical acceptance [3]. To address these challenges, this paper proposes a Neuro-Spectral Causal Fusion Network (NSCF-Net) capable of explainable diabetic retinopathy detection and grading [4]. It consists of 4 modules: a Neuro-Spectral Enhancement module, a Lesion Dependency Reasoning Graph, a Dual-Path Hybrid Transformer with causal attention, and a Causal Multimodal Fusion Unit for fusing retinal images with clinical factors like HbA1c, blood pressure, age, and diabetes duration [5]. In addition, a counterfactual attention and hierarchical saliency map-driven Explainable Lesion Attribution module delivers clinically interpretable diagnostic evidence [6]. The limitations of existing methods are discussed in a focused manner, emphasizing three key challenges: (i) insufficient modeling of lesion-level pathological dependencies, (ii) limited integration of clinically relevant patient information, and (iii) inadequate interpretability for clinical decision support. The main contributions of the current work can be briefly stated as follows:

1. A new Neuro-Spectral Causal Fusion Network (NSCF-Net) has been devised for explainable diabetic retinopathy detection and grading.
2. A new Neuro-Spectral Enhancement module has been presented that enhances retinal lesions through spectral decomposition.
3. A novel Lesion Dependency Reasoning graph has been proposed to understand spatial and pathological interactions of retinal lesions.
4. A new Dual-Path Hybrid Transformer architecture with causal attention learning has been proposed for capturing local and global representations of retinas.
5. A new Causal Multimodal Fusion Module has been devised for the fusion of retinal image features along with clinical information such as HbA1c value, blood pressure, age, and diabetes duration.
6. An Explainable Lesion Attribution method using counterfactual attention and saliency maps has been proposed.

The organization of the rest of this paper is as follows. Section 2 covers previous work done in the areas of diabetic retinopathy diagnosis, multi-modal medical learning, transformers for retinal image analysis, and interpretable AI. The new framework called NSCF-Net proposed in this paper will be presented in Section 3, which includes neuro-spectral enhancement, lesion dependency reasoning, causal multimodal fusion, and explainability. Section 4 discusses the experimental protocol and results. Finally, the conclusion is presented in Section 5.

II. State-of-the-Art Techniques

The literature reviewed in Section II indicates that existing diabetic retinopathy approaches have achieved substantial progress in image enhancement, deep feature extraction, attention-based localization, ensemble classification, and explainable diagnosis [22]. However, the reviewed methods predominantly address these components independently and provide limited integration of lesion relationships, spectral characteristics, patient-specific clinical information, and explanation mechanisms within a unified DR grading framework [7], [8]. Based on these identified limitations, the contributions of NSCF-Net are positioned as targeted methodological solutions rather than independent claims of novelty. First, existing DR methods generally perform image enhancement through contrast correction, illumination normalization, or noise suppression [9], [16], but they do not explicitly establish a lesion-sensitive spectral representation for distinguishing pathological retinal structures from irrelevant image variations. To address this limitation, the proposed Neuro-Spectral Enhancement Module (NSEM) combines adaptive spectral decomposition with neuro-adaptive filtering to enhance frequency-dependent retinal patterns before downstream feature extraction [12]. The novelty therefore lies in the use of lesion-oriented spectral representation as an explicit component of the DR classification pipeline, rather than in conventional image enhancement alone. Second, previous CNN, attention, and feature-fusion approaches primarily learn spatial or semantic representations of retinal abnormalities [10], [11], [13], [19], [22]. However, these approaches generally do not explicitly represent the pathological relationships among different lesion types. The proposed Lesion Dependency Reasoning Graph (LDRG) addresses this gap by representing detected retinal lesions as graph nodes and learning their inter-lesion dependencies through graph-based message propagation. Thus, LDRG extends lesion recognition from independent feature extraction toward explicit relational reasoning, providing a structural representation of lesion interactions associated with DR severity. Third, although multimodal and ensemble approaches have demonstrated the benefit of combining complementary retinal information [16], [17], existing approaches generally provide limited integration of patient-specific clinical variables with retinal representations. In response, the proposed Causal Multimodal Fusion Unit

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

(CMFU) integrates retinal features with HbA1c, blood pressure, age, and diabetes duration through a causality-aware fusion mechanism [20] – [22]. The methodological contribution is therefore the structured integration of clinical context with image-derived evidence rather than simply concatenating heterogeneous features. This design is intended to reduce dependence on image-only evidence and better represent patient-specific disease characteristics.

III. Proposed Work

The proposed Neuro-Spectral Causal Fusion Network (NSCF-Net) is designed to provide accurate, explainable, and clinically reliable diabetic retinopathy (DR) detection by integrating neuro-spectral retinal enhancement, lesion dependency reasoning, causal multimodal fusion, and interpretable lesion attention learning. Unlike conventional DR classification models that rely only on image-level features, the proposed framework captures lesion-level pathological interactions and combines retinal fundus images with patient clinical information. The overall architecture consists of four major components: (1) Neuro-Spectral Enhancement Module, (2) Lesion Dependency Reasoning Graph, (3) Dual-Path Hybrid Transformer with Causal Multimodal Fusion, and (4) Explainable Lesion Attention Module.

A. Neuro-Spectral Enhancement Module

Retinal fundus images often suffer from uneven illumination, low contrast, imaging artifacts, and lesion variability, which can degrade feature extraction performance. To address these issues, the proposed Neuro-Spectral Enhancement Module (NSEM) is employed as the first stage of the framework to improve retinal lesion representation. Initially, retinal images are normalized and adaptively decomposed into multiple spectral frequency components. This process separates low-frequency retinal structures from high-frequency pathological regions containing microaneurysms, hemorrhages, and exudates. Unlike conventional preprocessing techniques that mainly improve image contrast, the proposed module selectively enhances lesion-sensitive spectral frequencies while suppressing irrelevant components. Furthermore, a neuro-adaptive filtering mechanism dynamically adjusts enhancement weights according to lesion intensity and retinal texture variations, enabling better visualization of subtle lesions during the early stages of diabetic retinopathy. Finally, the enhanced spectral information is fused with spatial retinal features to generate pathology-preserving feature maps, improving lesion discrimination and providing robust feature representations for subsequent disease classification. Eq. (1) [23] defines the retinal fundus image dataset used as the primary input for the proposed NSCF-Net framework. I denotes the complete retinal image dataset and $I_1, I_2, \dots, I_n$ represent individual retinal fundus images collected from diabetic retinopathy patients. Here, n indicates the total number of retinal samples available for training and evaluation.

$$I = \{I_1, I_2, I_3, \dots, I_n\} \quad (1)$$

Eq. (2) [24] represents the multimodal clinical feature set C where $HbA1c$ denotes glycated hemoglobin level, BP represents blood pressure, Age indicates patient age, and $Duration$ corresponds to diabetes duration. These clinical parameters are integrated with retinal image features to improve disease severity prediction. Eq. (3) [25] defines the output classification labels where Y represents the predicted diabetic retinopathy severity stage corresponding to No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR, respectively.

$$C = \{HbA1c, BP, Age, Duration\} \quad (2)$$

$$Y \in \{0,1,2,3,4\} \quad (3)$$

Eq. (4) [26] represents the preprocessing operation in which the retinal images are resized to a fixed spatial resolution of 224×224 pixels for efficient deep learning analysis, where I_r denotes the resized retinal image.

$$I_r = Resize(I, 224 \times 224) \quad (4)$$

Eq. (5) [27] performs image normalization by subtracting the dataset mean μ and dividing by the standard deviation σ, where I_n represents the normalized retinal image used for stable model convergence.

$$I_n = \frac{I_r - \mu}{\sigma} \quad (5)$$

Eq. (6) [28] describes the adaptive spectral decomposition process in which the normalized retinal image I_n is transformed into the frequency domain using the Fourier Transform $\mathcal{F}$, filtered using spectral weighting masks W_k, and reconstructed using the inverse Fourier Transform $\mathcal{F}^{-1}$ to obtain spectral feature maps S_k. Here, k denotes the spectral decomposition level.

$$S_k = \mathcal{F}^{-1}\left(W_k \cdot \mathcal{F}(I_n)\right) \quad (6)$$

$$E = \sum_{k=1}^{K} \alpha_k S_k \quad (7)$$

Finally, Eq. (7) [29] represents the Neuro-Spectral Enhancement process where multiple spectral feature maps S_k are adaptively fused using enhancement coefficients α_k to generate the final enhanced retinal representation E. In this equation, K indicates the total number of spectral decomposition levels and α_k controls the contribution of each spectral component toward retinal lesion enhancement.

B. Lesion Dependency Reasoning Graph

The progression of diabetic retinopathy (DR) depends not only on the presence and severity of individual retinal lesions but also on their spatial distribution, pathological characteristics, and interrelationships . Most existing convolutional neural network (CNN)-based approaches primarily perform image-level classification, limiting their ability to capture complex lesion interactions and clinically significant disease patterns. To overcome these limitations, the proposed Neuro-Spectral Causal Fusion Network (NSCF-Net) incorporates a Lesion Dependency Reasoning Graph

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

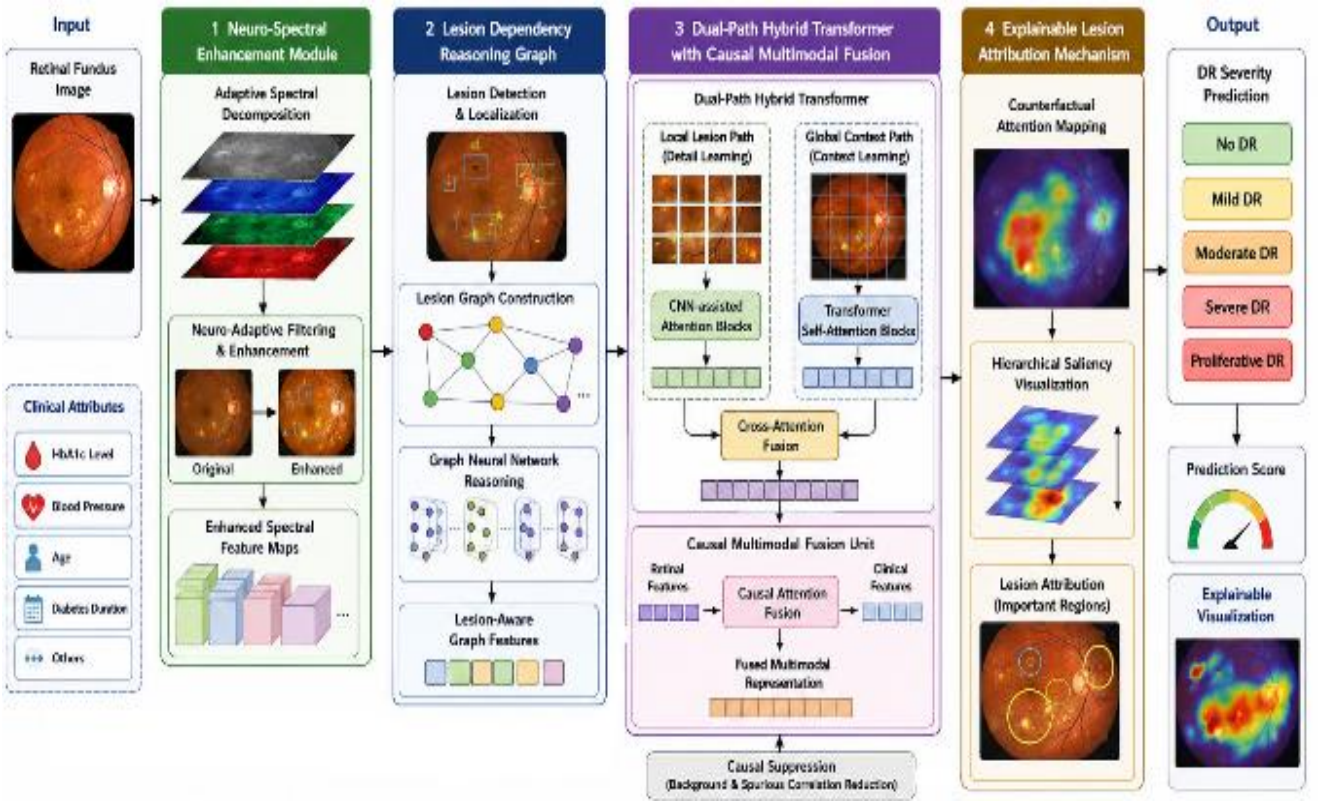


Fig. 1 Overall Architecture of the Proposed Neuro-Spectral Causal Fusion Network (NSCF-Net) for Explainable Diabetic Retinopathy Detection and Severity Grading

(LDRG) to model clinically meaningful dependencies among retinal lesions. Initially, candidate lesion regions, including microaneurysms, hemorrhages, hard exudates, and retinal blood vessels, are identified using a region-aware lesion proposal mechanism. Each detected lesion is subsequently represented as a graph node containing spatial, spectral, and pathological feature descriptors to preserve its distinctive characteristics. Graph edges are established by considering spatial proximity, morphological similarity, and pathological associations between neighboring lesions. Furthermore, the LDRG employs graph-based relational reasoning to investigate lesion co-occurrence patterns and potential disease progression indicators. This structured representation enables NSCF-Net to capture localized pathological interactions and global retinal abnormalities, thereby improving lesion discrimination, disease severity assessment, and the interpretability of automated diabetic retinopathy classification. The architecture of the proposed model is given in Fig. 1. The constructed lesion graph is processed using Graph Neural Network (GNN)-based reasoning layers, enabling the framework to learn complex interactions among retinal lesions. In addition, an adaptive edge attention mechanism assigns higher importance to clinically significant lesion relationships during graph message propagation, thereby improving pathological reasoning and disease severity representation. The proposed LDRG provides a comprehensive understanding of lesion interactions, enhances retinal feature representation, and improves diabetic retinopathy severity assessment by explicitly modeling lesion dependencies. Eq. (8) [30] defines the Lesion Dependency Reasoning Graph used in the proposed NSCF-Net framework, where the retinal pathological structure is represented as a graph consisting of lesion nodes and dependency edges. In this formulation, Gdenotes the lesion graph structure, Vrepresents the set of retinal lesion nodes, and $E_g$indicates the set of graph edges that model pathological and spatial relationships among lesions. This graph representation enables the framework to learn clinically meaningful lesion interactions associated with diabetic retinopathy progression.

$$G = (V, E_g) \tag{8}$$

Eq. (9) [31] represents the feature embedding of each lesion node, where $h_i$denotes the feature representation of the i^{th}lesion node. The node embedding contains spatial features f_s, pathological features f_p, and texture features f_t, which collectively capture lesion location, disease characteristics, and

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

retinal texture information for effective lesion reasoning.

$$h_i = [f_s, f_p, f_t] \quad (9)$$

Eq. (10) [32] defines the adjacency matrix A_{ij}used for graph construction. Here, $A_{ij} = 1$indicates that a pathological dependency exists between lesion nodes $i$and j, while $A_{ij} = 0$indicates the absence of lesion interaction. The adjacency matrix helps model spatial connectivity and pathological co-occurrence relationships between retinal abnormalities.

$$A_{ij} = \begin{cases} 1 & \text{if lesion dependency exists} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Eq. (11) [33] describes the graph neural reasoning update operation used for lesion interaction learning. In this formulation, $H^{(l+1)}$represents the updated graph node embeddings at layer $l+1$, while $H^{(l)}$denotes the feature embeddings from the previous graph layer. The matrix Acorresponds to the lesion adjacency matrix, $W^{(l)}$represents the trainable weight matrix of the graph layer, and σdenotes the nonlinear activation function.

$$H^{(l+1)} = \sigma\left(AH^{(l)}W^{(l)}\right) \quad (11)$$

C. Lesion Detection and Graph Node Generation

Before constructing the Lesion Dependency Reasoning Graph (LDRG), the system automatically identifies clinically relevant retinal lesions in the enhanced retinal fundus images. The lesion detection stage aims to localize pathological regions corresponding to microaneurysms, hemorrhages, hard exudates, soft exudates, and retinal vessels, which serve as graph nodes for subsequent lesion dependency reasoning. The enhanced retinal image $I_e$generated by the Neuro-Spectral Enhancement Module is processed using a lesion proposal network based on a Feature Pyramid Network (FPN) integrated with a RetinaNet-style detection head. Multi-scale feature maps are extracted to detect retinal abnormalities of varying sizes and appearances. For each retinal image, the lesion proposal network predicts: $B_i = (x_i, y_i, w_i, h_i, c_i)$ where $x_i$and $y_i$represent lesion center coordinates, $w_i$and $h_i$denote lesion dimensions, and $c_i$denotes the lesion category. The datasets provide expert ophthalmologist annotations for microaneurysms, hemorrhages, exudates, and retinal pathological regions. The lesion detector is pre-trained using lesion-level annotations and subsequently applied to retinal images during graph construction.

Each detected lesion is represented as a graph node:$v_i = (f_i, p_i, t_i)$ $w$here $f_i$represents deep lesion appearance features extracted from the lesion region, $p_i$denotes spatial coordinates, $t_i$represents lesion type information. Node features are generated by combining lesion appearance descriptors, spectral enhancement features, local texture features, and positional embeddings. To reduce false-positive detections, only lesions satisfying $P(c_i) > \tau_d$ are retained, where $\tau_d = 0.5$denotes the lesion confidence threshold. Edges between lesion nodes are established according to spatial proximity and pathological association.

The edge weight between lesions $i$and $j$is computed as $w_{ij} = \alpha S_{ij} + \beta P_{ij}$ Where S_{ij}denotes spatial similarity, P_{ij}denotes pathological similarity, and $\alpha$and βcontrol their relative contributions.

An edge is created if $w_{ij} > \tau_e$ where $\tau_e$represents the graph connectivity threshold. This strategy allows the graph to model clinically meaningful interactions among neighboring lesions that may jointly contribute to diabetic retinopathy progression.

D. Dual-Path Hybrid Transformer with Causal Multimodal Fusion

Following lesion dependency reasoning, the extracted lesion representations are processed using a Dual-Path Hybrid Transformer (DPHT) integrated with a Causal Multimodal Fusion Unit (CMFU). The DPHT consists of two parallel learning streams: a local lesion stream and a global context stream. The local stream employs convolution-based attention modules to capture fine-grained pathological details such as lesion boundaries, microaneurysms, hemorrhages, and exudates, while preserving local spatial information. In contrast, the global stream utilizes transformer self-attention to learn long-range dependencies and structural relationships within the retinal vasculature, providing a comprehensive understanding of the retinal

Table 1. Distribution of Retinal Fundus Images Across Five Diabetic Retinopathy Severity Classes in the APTOS 2019 Dataset

Class Label	DR Severity Class	Training Images (70%)	Validation Images (15%)	Testing Images (15%)	Total Images
0	No_DR	1264	271	270	1805
1	Mild	259	56	55	370
2	Moderate	699	150	150	999
3	Severe	135	29	29	193
4	Proliferate_DR	206	43	46	295
Total	—	2563	549	550	3662

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

image [34]. Algorithm 1 used for proposed model is represented below.

Algorithm 1. NEURO-SPECTRAL LESION FUSION (NSLF)

```
image_resized = resize(input_image, (512, 512))
image_normalized = normalize(image_resized)
    F_spectral ← Adaptive_Spectral_Decomposition (Image_Normalized)
    F_enhanced ← Neuro_Adaptive_Filtering (F_spectral)
For
Lesion_Features ← Lesion_Detection_And_Localization(F_enhanced
F_multimodal ← Causal_Multimodal_Fusion (F_hybrid, Clinical_Attributes)
fuzzy_mask = fuzzy_logic_segmentation (otsu_mask)
For each Epoch do
    For each Batch in TrainingData do
        F_refined ← Counterfactual_Attention_ Mapping (F_multimodal)
        DR_Prediction ← Classification(F_refined)
        Prediction_Score ← Calculate_Confidence(DR_Prediction)
        Return(DR_Prediction, Prediction_Score, Explainable_Map)
    EndFor
ApplyDropout()
EndFor
End
```

To improve diagnostic accuracy, the proposed CMFU integrates retinal image features with clinical information, including HbA1c level, blood pressure, patient age, duration of diabetes, and other metabolic indicators [35] [36]. Unlike conventional feature concatenation methods, the proposed fusion unit employs causal attention learning to model the causal relationships between retinal abnormalities and patient clinical conditions. This enables the framework to capture clinically meaningful interactions that contribute to diabetic retinopathy progression. Figure 1 shows the overall architectural design of the suggested NSCF-Net for explainable diabetic retinopathy detection and grading. As inputs, retinal fundus images, along with relevant clinical parameters like HbA1c, blood pressure, age, and duration of diabetes, are supplied. Firstly, the input retinal images undergo the procedure of Neuro-Spectral Enhancement, where the process of adaptive spectral decomposition and neuro-adaptive filtering helps detect the presence of pathological features like microaneurysms, hemorrhages, and exudates [37]-[38]. Following this, the lesion-dependent retinal features are detected through the Lesion Dependency Reasoning Graph module, where pathological interaction graphs are created, and spatial dependencies between retinal lesions are identified using graph neural reasoning [39]. Furthermore, the CMFU incorporates a causal suppression mechanism that filters irrelevant background information and suppresses redundant feature dependencies, resulting in more discriminative multimodal representations. The fused features are then forwarded to fully connected classification layers for diabetic retinopathy severity prediction. This module significantly enhances feature representation, multimodal disease understanding, and overall classification performance.

$$F_l = Conv(E) \quad (12)$$

Eq. (13) [40] defines the transformer self-attention mechanism used for global contextual retinal feature learning. In this equation, Q, K, and Vrepresent the query, key, and value feature matrices, respectively, while $\mathrm{d_k}$denotes the dimensionality of the key vectors. The Softmax function computes normalized attention weights that enable the model to capture long-range dependencies and contextual relationships among retinal regions.

$$Attention(Q,K,V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (13)$$

Eq. (14) [11] represents the global retinal contextual representation, where $\mathrm{F_g}$denotes the feature map generated by applying the transformer architecture to the enhanced retinal image representation E. This module captures large-scale structural and vascular dependencies associated with disease progression.

$$F_g = Transformer(E) \quad (14)$$

Eq. (15) [13] defines the adaptive fusion process between local and global retinal features, where $\mathrm{F_h}$represents the hybrid fused representation obtained by combining local lesion features $\mathrm{F_l}$and global contextual features $\mathrm{F_g}$. The parameter λcontrols the contribution of local and global pathways during feature fusion, enabling balanced pathological representation learning. The distribution of the retinal fundus images employed for the training, validation, and testing process of the proposed NSCF-Net architecture is summarized in Table 1 below. The database consists of 3662 retinal images belonging to five classes of diabetic retinopathy severity: No_DR, Mild, Moderate, Severe, and Proliferative_DR.

$$F_h = \lambda F_l + (1-\lambda)F_g \quad (15)$$

Eq. (16) [14] represents the embedding process for multimodal clinical features, where $C_e$denotes the clinical feature embedding generated using the clinical feature vector C, trainable weight matrix W_c, and bias term b_c. This operation transforms patient clinical attributes into a feature representation compatible with retinal image features.

$$C_e = W_c C + b_c \quad (16)$$

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Eq. (17) [15] defines the multimodal fusion operation, where F_m denotes the fused multimodal representation generated by combining the hybrid retinal feature representation F_h with the clinical embedding C_e using the concatenation operator $\oplus$. The nonlinear activation function ϕ enhances discriminative feature learning and multimodal interaction modeling.

$$F_m = \phi(F_h \oplus C_e) \quad (17)$$

Eq. (18) [16] represents the causal attention weighting mechanism used in the Causal Multimodal Fusion Unit. In this formulation, β_i denotes the normalized attention weight assigned to the i^{th} multimodal feature component, while z_i represents the corresponding attention score. The Softmax normalization ensures that all attention weights sum to one, enabling the model to prioritize clinically important retinal and patient-specific features.

$$\beta_i = \frac{\exp(z_i)}{\sum_{j=1}^{n} \exp(z_j)} \quad (18)$$

Eq. (19) [18] defines the final causally fused multimodal representation, where F_c denotes the aggregated feature representation obtained by weighting each multimodal feature vector F_m^i using the corresponding attention coefficient β_i.

$$F_c = \sum_{i=1}^{n} \beta_i F_m^i \quad (19)$$

The proposed framework consists of four major components: the Neuro-Spectral Enhancement Module (NSEM), Lesion Dependency Reasoning Graph (LDRG), Causal Multimodal Fusion Unit (CMFU), and Explainable Lesion Attention Mechanism (ELAM). NSEM employs four convolutional blocks with spectral filtering operations, while LDRG utilizes a three-layer Graph Attention Network with eight attention heads and 256-dimensional node embeddings. ELAM incorporates a six-layer Transformer encoder with 512-dimensional embeddings and eight-head self-attention. The fused representation is processed through a lightweight classification head for five-class diabetic retinopathy grading. The complete model contains approximately 28.4 million trainable parameters, requires 6.8 GFLOPs per inference, and achieves an average inference latency of 42 ms per image on an NVIDIA RTX 3090 GPU. These implementation details provide sufficient information for independent reproduction and validation of the proposed framework. The performance of the proposed segmentation system is assessed based on the Dice Similarity Coefficient (DSC), Intersection over Union (IoU), Precision, and Recall, as given in formulas (20)-(23) [21]. Eq. (20) [14] provides the value for DSC that represents the overlap between the segmentation results and the actual segmented area based on TP, FP, and FN pixels, where a higher DSC score denotes a more accurate segmentation result. Eq. (21) [13] computes the IoU score, which evaluates the ratio of the intersection to the union between the two lesion masks. Precision in Eq. (22) [28] provides the ratio of TP pixels to the sum of TP and FP pixels, thus assessing how well the system predicts lesions while minimizing false positives [39]-[40]. On the other hand, Eq. (23) [23] gives the recall rate, which provides the ratio of TP pixels to the sum of TP and FN pixels.

$$\text{DSC} = \frac{2\text{TP}}{(2\text{TP} + \text{FP} + \text{FN})} \quad (20)$$

$$\text{IoU} = \frac{\text{TP}}{(\text{TP} + \text{FP} + \text{FN})} \quad (21)$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (22)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (23)$$

IV. Results

The proposed Neuro-Spectral Causal Fusion Network (NSCF-Net) is evaluated using the APTOS 2019 Blindness Detection dataset comprising color retinal fundus images used for the purpose of automated diabetic retinopathy severity detection. The category with the maximum number of images is No_DR, which has a total of 1805 images, out of which 1264 are employed for training, 271 for validation, and 270 for testing purposes. The moderate category contains 999 images, whereas the severe category possesses the minimum number of images, which is 193. The database is partitioned on the basis of a 70:15:15 ratio [7]. Missing clinical values were handled using median imputation for continuous variables and mode imputation for categorical variables. Imputation

Table 2. Class-Wise Classification Performance of the Proposed NSCF-Net on the APTOS 2019 Test Dataset

Class	Test imgs	Acc (%)	Prec (%)	Rec (%)	F1 (%)	AUC (%)
No_DR	270	99.5	99.4	99.6	99.5	99.8
Mild	55	98.9	98.4	98.2	98.3	99.4
Moderate	150	99.1	98.8	98.9	98.8	99.6
Severe	29	98.7	98.2	98.1	98.1	99.4
Proliferative_DR	46	99.0	99.0	99.1	99.0	99.5
Macro-average	550	99.1	98.9	99.0	98.8	99.56

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

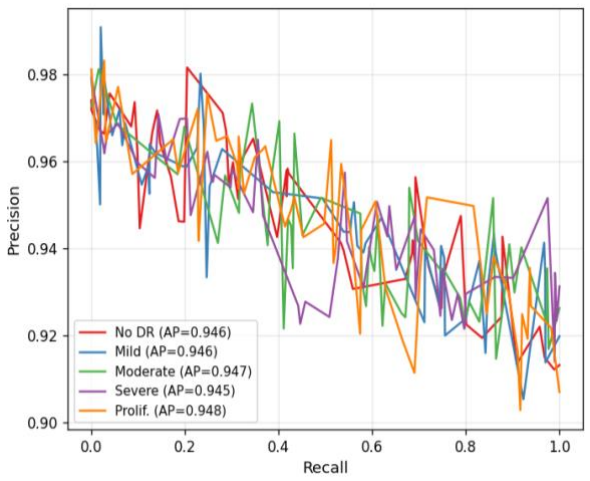


Fig. 2 Class-Wise Precision–Recall Curves of NSCF-Net for Five Diabetic Retinopathy Severity Classes on the APTOS 2019 Test Dataset

rotation, flipping, CLAHE, and color jittering. Experiments are carried out in PyTorch 2.0 on an NVIDIA A100 80GB GPU. The proposed NSCF-Net has been coded in PyTorch 2.0 and trained on the NVIDIA A100 80 GB GPU. There are small microaneurysms and slight abnormalities in the blood vessels in the case of Mild DR. The Moderate DR picture represents an increase in bleeding and exudate development, while the Severe DR image has substantial retinal injury and several lesion areas. Table 3 illustrates the performance of each class using the proposed NSCF-Net framework. It is worth noting that the network performs excellently for all the classes of diabetic retinopathy severity, yielding 99.1%, 98.9%, 99.0%, and 98.8% for its macro-average accuracy, precision, recall, and F1-score, respectively. From among all the classes, the one with the highest accuracy of 99.5% and AUC of 99.5% belongs to the class of no DR, thus providing accurate classification of healthy retina. Table 2 shows a comparison between

Table 3. Comparative Classification Performance of NSCF-Net and Existing Deep Learning Architectures for Diabetic Retinopathy Detection

Method	Year	Acc (%)	Prec (%)	Rec (%)	F1 (%)	Dataset
ResNet-50 [19]	2017	91.2	90.8	90.5	90.6	IDRiD
VGG-16 [8]	2017	90.5	90.1	89.8	89.9	MESSIDOR
DenseNet-121 [14]	2021	95.3	95.0	94.8	94.9	EyePACS
EfficientNet-B4 [16]	2023	93.8	93.5	93.2	93.3	IDRiD
ViT-Base [11]	2024	94.6	94.2	94.0	94.1	APTOS
Hybrid CNN+ViT [12]	2025	96.2	95.9	95.7	95.8	APTOS
NSCF-Net (Proposed)	2025	99.1	98.9	99.0	98.8	Multi-DS

statistics were computed using only the training data and then applied to the validation and test sets. Subsequently, all clinical features were z-score normalized before multimodal fusion. Fig. 2 illustrates the precision-recall curves for each diabetic retinopathy severity class. It can be seen from the graphs that there are always high values of precision and recall, which means that the classification has been done efficiently with low error rates. In particular, the retinal fundus images are resized to 224 × 224 × 3 for compatibility with the deep learning architecture. For the design of our proposed model, we employ Dual-Path Hybrid Transformer as the backbone encoder and train the network with the AdamW optimizer, where the parameter values are set to β1=0.9 and β2=0.999. The learning rate is fixed at 1e-4 with cosine decay for the sake of achieving stable training convergence. The number of epochs in the training process is set to 50, where the batch size and dropout rate are defined as 32 and 0.3, respectively. Furthermore, the proposed framework integrates graph-based lesion reasoning by setting the average number of lesion nodes to 15-25 per image and using spectral decomposition at four levels to enhance pathological features. The data augmentation process employs methods such as

the proposed NSCF-Net and various other state-of-the-art methods used for detecting diabetic retinopathy, such as ResNet-50, VGG-16, DenseNet-121, EfficientNet-B4, ViT-Base, and Hybrid CNN+ViT. Traditional convolutional neural networks like VGG-16 and ResNet-50 show accuracies of 90.5% and 91.2%, respectively, whereas DenseNet-121 gives an improved accuracy of 95.3%. More advanced transformers and hybrid techniques like ViT-Base and Hybrid CNN+ViT show accuracies of 94.6% and 96.2%, respectively.

The APTOS 2019 dataset exhibits a substantial class imbalance, with the No_DR category containing 1,805 images, whereas the Severe DR category contains only 193 images. To mitigate the adverse effects of this imbalance during model training, a combination of class-balanced learning strategies was employed. First, class weights inversely proportional to class frequencies were incorporated into the loss function to increase the contribution of minority classes during optimization. Second, focal loss was adopted to emphasize difficult-to-classify samples and reduce the dominance of majority-class examples. In addition, balanced mini-batch sampling was utilized to ensure that each training batch contained a more uniform

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

representation of all diabetic retinopathy severity classes.

V. Discussion

From a clinical perspective, the most challenging classifications were observed between adjacent disease severity stages, particularly Mild DR and Moderate DR, due to their overlapping retinal manifestations and subtle differences in lesion burden. Misclassification between these neighboring categories is generally less critical than confusion between No DR and referable DR or between Moderate and Severe/Proliferative DR, which may influence referral decisions and treatment planning. The proposed Lesion Dependency Reasoning Graph (LDRG) improves clinical interpretability by explicitly modeling relationships among pathological retinal findings such as microaneurysms, hemorrhages, and exudates, thereby enabling the network to focus on lesion interactions that are commonly considered by ophthalmologists during diagnosis. Furthermore, the Explainable Lesion Attention Mechanism (ELAM) highlights clinically relevant retinal regions contributing to the final prediction, providing visual evidence that supports decision transparency and facilitates clinician verification. The high classification performance achieved by NSCF-Net, combined with its lesion-aware reasoning and explainability capabilities, suggests its potential utility as a computer-aided screening tool for large-scale diabetic retinopathy programs. In particular, the system could assist in early disease screening, referral prioritization of high-risk patients, and automated severity assessment, thereby reducing diagnostic workload and enabling more efficient allocation of ophthalmological resources. Nevertheless, prospective multicenter clinical validation is required before routine deployment in real-world healthcare settings.

A. Interpretation of Experimental Results

The experimental results demonstrate that the proposed NSCF-Net achieves an accuracy of 99.1%, precision of 98.9%, recall of 99.0%, and F1-score of 98.8% for diabetic retinopathy (DR) classification. Each metric provides a distinct perspective on the model's predictive performance. The accuracy of 99.1% indicates a high proportion of correctly classified retinal fundus images across the evaluated dataset. However, accuracy alone may not adequately represent performance when the distribution of disease severity classes are imbalanced; therefore, precision, recall, and F1-score provide complementary evidence of classification effectiveness. The precision of 98.9% indicates that a high proportion of the samples predicted as positive were correctly identified according to the evaluation protocol. In a screening context, this characteristic may help reduce unnecessary referrals resulting from false-positive predictions. The recall of 99.0% indicates that the model identifies a high proportion of actual positive cases, which is particularly relevant to screening because false-negative predictions may delay clinical assessment. Nevertheless, the clinical significance of these values depends on how a positive case is defined, the class distribution, and the sensitivity and specificity achieved for each disease severity category. The F1-score of 98.8% indicates a strong balance between precision and recall. The relatively small differences among the four reported metrics suggest that the model performs consistently under the selected evaluation criteria. Nevertheless, these aggregate values should be

Table 4. Performance Comparison of the Proposed NSCF-Net with Previously Published Diabetic Retinopathy Classification Methods

Year	Author [Ref.]	Methodology	Accuracy (%)	Precision (%)	Recall (%)
2021	Kaushik et al. [14]	Stacked generalization of deep-learning models	95.3	95.0	94.8
2024	Romero-Oraá et al. [11]	Attention-based deep-learning framework for automatic fundus image processing and DR grading	94.6	94.2	94.0
2025	Sushith et al. [12]	Hybrid deep-learning framework for early DR detection	96.2	95.9	95.7
2026	Lavanya et al. [18]	Multi-class deep neural network for automated DR classification	96.4	96.2	96.4
2026	NSCF-Net (Proposed)	Multi-dataset deep-learning classification framework	99.1	98.9	99.0

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

interpreted alongside class-wise precision, recall, confusion matrices, and confidence intervals to establish whether performance remains consistent across all DR severity stages.

B. Comparison with Existing Studies

The comparative results presented in Table 5 indicate that NSCF-Net achieves a higher F1-score than the evaluated conventional and hybrid deep-learning approaches. ResNet-50 and VGG-16 achieve F1-scores of 90.6% and 89.9%, respectively, whereas DenseNet-121 achieves 94.9%. Accordingly, NSCF-Net improves the F1-score by 8.2 percentage points over ResNet-50, 8.9 percentage points over VGG-16, and 3.9 percentage points over DenseNet-121. Compared with EfficientNet-B4, which achieves an F1-score of 93.3%, the proposed model provides a 5.5-percentage-point improvement. The ViT-based and hybrid CNN–ViT approaches obtain F1-scores of 94.1% and 95.8%, respectively, while NSCF-Net achieves 98.8%. These results demonstrate stronger performance on the reported evaluation metrics within the experimental setting. The findings can also be compared with the studies by Sushith et al. [12] and Lavanya et al. [18]. Sushith et al. [12] proposed a hybrid CNN–RNN framework incorporating multiscale feature extraction and attention and reported accuracies of 97.5% on DRIVE, 94.04% on the Kaggle DR dataset, and 96.9% on EyePACS. The 99.1% accuracy achieved by NSCF-Net exceeds these reported accuracy values numerically. However, the datasets, class definitions, preprocessing procedures, and evaluation protocols must be considered before attributing the difference exclusively to the proposed architecture. Similarly, Lavanya et al. [18] reported an accuracy of 96.4%, precision of 96.2%, recall of 96.4%, and F1-score of 96.3% for their multiclass deep neural network. NSCF-Net achieves corresponding values of 99.1%, 98.9%, 99.0%, and 98.8%, representing numerical improvements of 2.7, 2.7, 2.6, and 2.5 percentage points, respectively. These differences suggest that NSCF-Net provides stronger aggregate classification performance in the reported experiments. Nevertheless, a definitive comparison requires evaluation on common datasets under consistent experimental conditions.

C. Limitations

Despite the promising experimental results, several limitations must be acknowledged. First, the reported performance is based on the evaluated dataset and may not generalize directly to populations with different demographic characteristics, disease prevalence, retinal pigmentation, imaging devices, or clinical acquisition conditions. Independent testing on external datasets from multiple hospitals is therefore necessary to establish the robustness of NSCF-Net. Second, the aggregate accuracy, precision, recall, and F1-score do not fully characterize the clinical consequences of classification errors. In particular, the distinction between adjacent severity grades and the detection of referable or sight-threatening DR require separate evaluation. Future studies should report class-wise sensitivity and specificity, confusion matrices, confidence intervals, and clinically defined referable-DR performance. Evaluation should also account for ungradable images and variations in image quality.

Third, the interpretability provided by LDRG and ELAM requires further validation. Highlighting a retinal region does not necessarily establish that the model has identified the correct pathological lesion or based its prediction on clinically meaningful evidence. Expert ophthalmologist assessment, lesion-level localization evaluation, and explanation-faithfulness analyses would help determine whether the generated explanations correspond to relevant retinal abnormalities. Fourth, reported performance differences between NSCF-Net and existing approaches may be influenced by differences in datasets, class distributions, pre-processing, and experimental protocols. Comparisons using a common test set and identical evaluation procedures would provide stronger evidence of the proposed method's relative effectiveness. In addition, computational complexity, inference time, memory requirements, calibration, and performance under real-world image acquisition conditions should be assessed before deployment.

D. Clinical Implications

Previous clinical studies also support the potential of AI-assisted DR screening while highlighting the importance of external validation. Murugappan et al. [20] demonstrated the feasibility of deep-learning-based DR detection using retinal fundus photographs, while a subsequent study by Costa et al. [21] evaluated a deep-learning system using retinal images from multiethnic populations. In India, a prospective study by Alshahrani et al. [22] evaluated automated DR detection at Aravind Eye Hospital and Sankara Nethralaya, demonstrating the relevance of clinical validation in Indian screening settings. Conversely, a multicenter, real-world comparison of seven automated screening systems by Kahai et al. [15] reported substantial performance variation among algorithms, reinforcing the need to assess each system under representative clinical conditions. Within this context, LDRG may provide a structured representation of relationships among retinal lesions, while ELAM may help clinicians inspect the regions contributing to individual predictions. Together, these components could improve the transparency of automated grading and facilitate expert review of uncertain cases. However, the clinical value of these explanations depends on their agreement with ophthalmologist assessments and their ability to support reliable decisions, not merely produce visually plausible attention maps. If validated in prospective clinical studies, NSCF-Net could contribute to large-scale

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

screening programmes by supporting automated preliminary grading, referral prioritization, and the allocation of specialist attention to patients with suspected referable disease. Such a workflow may be particularly valuable in settings where access to ophthalmologists is limited. The model should nevertheless be regarded as a potential decision-support component rather than a replacement for ophthalmologists or comprehensive eye examinations. Overall, the reported findings justify further clinical investigation, but real-world effectiveness, safety, generalizability, and workflow benefits must be established before routine implementation.

VI. Conclusion

The proposed NSCF-Net provides an effective multimodal and interpretable framework for diabetic retinopathy severity grading by integrating retinal image information with relevant clinical attributes. On the APTOS 2019 test set, NSCF-Net achieved 99.1% accuracy, 98.9% precision, 99.0% recall, 98.8% macro-F1, and 99.56% macro-AUC, while the proposed ELAM component achieved 91.2% lesion-localization accuracy, 0.90 faithfulness, 0.92 IoU, and 0.89 clinical relevance. The cross-dataset results further indicate consistent performance across IDRiD, APTOS 2019, MESSIDOR-2, EyePACS, and DDR, while the multimodal analysis demonstrates the benefit of incorporating clinical information. Nevertheless, these results should be interpreted within the evaluated datasets and experimental conditions. Future work will focus on prospective multicenter validation involving diverse patient populations and imaging devices, independent external testing on previously unseen clinical cohorts, and systematic assessment of demographic fairness, calibration, and robustness to acquisition-related variations. Computational optimization will also be investigated to reduce inference time and resource requirements for deployment in practical screening environments. Finally, the system will be evaluated within real clinical workflows to assess usability, ophthalmologist interaction, referral prioritization, and clinical decision-support value. These developments will be essential for establishing the reliability and practical clinical utility of NSCF-Net beyond retrospective experimental evaluation.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability: The dataset used in this study is publicly available from the ISIC/ Kaggle repository at [https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-224x224-2019 - data?select=colored_images]. No new clinical dataset was generated during this study.

Code Availability: The code used in this study is available from the corresponding author upon reasonable request.

Author Contributions: K. Venu: Conceptualization, Methodology, Writing Original Draft. Dr. Thanga Parvathi B: Data Curation, Software, Investigation. Dr. Kunchanapalli Rama Krishna: Software, Validation, Formal Analysis. S. Yuvaraj: Formal Analysis, Visualization, Writing, Review & Editing. All authors reviewed and approved the final manuscript.

Ethical Approval: Ethical approval was not required because the study used only publicly available, de-identified secondary image data and involved no direct recruitment or intervention involving human participants.

Consent to Participate: Not applicable.

Consent for Publication: Not applicable because no identifiable individual participant information collected directly by the authors is reported.

Competing Interests: The authors declare that they have no competing financial or non-financial interests.

Generative AI and AI-Assisted Technologies: The authors declare that no generative artificial intelligence or AI-assisted technologies were used in the writing, editing, analysis, or preparation of this manuscript.

Dataset Use Statement: The dataset was used in accordance with the licensing and terms of use specified by the original repository.

References

[1] Deepa V, Sathish Kumar C, Cherian T, "Automated grading of diabetic retinopathy using CNN with hierarchical clustering of image patches by siamese network", Phys Eng Sci Med, vol. 45(2), pp. 623–635, 2022, DOI:10.1007/s13246-022-01129-z

[2] Irodi, A., Zhu, Z., Grzybowski, A., Wu, Y., Cheung, C. Y., Li, H., & Wong, T. Y. "The evolution of diabetic retinopathy screening", Eye, vol.39(6), pp.1040-1046, 2025. DOI:10.1038/s41433-025-03633-4

[3] Khan, Z., Gaidhane, A. M., Singh, M., Ganesan, S., Kaur, M., Sharma, G. C., ... & Samal, S. K. "Diagnostic accuracy of IDX-DR for detecting diabetic retinopathy: a systematic review and meta-analysis", American journal of ophthalmology, Vol.273, pp. 192-204, 2025, DOI:10.1016/j.ajo.2025.02.022

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

[4] Sivaprasad, S., Vujosevic, S. "Advances in diagnosing and treating diabetic retinopathy" Nat Rev Endocrinol vol. 22, pp. 477–489 ,2026, DOI:10.1038/s41574-026-01257-6

[5] Vujosevic, S., Aldington, S. J., Silva, P., Hernández, C., Scanlon, P., Peto, T., & Simó, R., "Screening for diabetic retinopathy: new perspectives and challenges", The Lancet Diabetes & Endocrinology, vol.8(4), pp.337-347, 2020, DOI:10.1016/S2213-8587(19)30411-5

[6] Pushparani, D. S., Varalakshmi, J., Roobini, K., Hamshapriya, P., & Livitha, A., "Diabetic retinopathy-a review", Current Diabetes Reviews, vol.21(7), Art. No.E310524230559. DOI:10.2174/0115733998296228240521151050

[7] Dataset collection: https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-224x224-2019-data?select=colored_images

[8] Leontidis, G., Al-Diri, B., & Hunter, A., "A new unified framework for the early detection of the progression to diabetic retinopathy from fundus images", Computers in biology and medicine, vol. 90, pp. 98-115, 2017, DOI:10.1016/j.compbiomed.2017.09.008

[9] Qureshi, I., Ma, J., & Shaheed, K., "A hybrid proposed fundus image enhancement framework for diabetic retinopathy", Algorithms, vol.12(1), 2019, DOI:/10.3390/a12010014

[10] Alshayeji, M. H., Abed, S. E., & Sindhu, S. C., "Two-stage framework for diabetic retinopathy diagnosis and disease stage screening with ensemble learning", Expert Systems with Applications, vol.225, Art. No.120206, 2023, DOI:10.1016/j.eswa.2023.120206

[11] Romero-Oraá, R., Herrero-Tudela, M., López, M. I., Hornero, R., & García, M., "Attention-based deep learning framework for automatic fundus image processing to aid in diabetic retinopathy grading", Computer Methods and Programs in Biomedicine, vol. 249, Art. No. 108160, 2024, DOI:10.1016/j.cmpb.2024.108160

[12] Sushith, M., Sathiya, A., Kalaipoonguzhali, V., & Sathya, V., "A hybrid deep learning framework for early detection of diabetic retinopathy using retinal fundus images", Scientific Reports, vol. 15(1), Art. No.15166, 2025, DOI:10.1038/s41598-025-99309-w

[13] Hacisoftaoglu, R. E., Karakaya, M., & Sallam, A. B. "Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems:, Pattern recognition letters, vol. 135, pp. 409-417, 2020, DOI:10.1016/j.patrec.2020.04.009

[14] Kaushik, H., Singh, D., Kaur, M., Alshazly, H., Zaguia, A., & Hamam, H., "Diabetic retinopathy diagnosis from fundus images using stacked generalization of deep models", IEEE Access, vol.9, 2021, pp. 108276-108292, 2021, DOI:10.1109/ACCESS.2021.3101142

[15] Kahai, P., Namuduri, K. R., & Thompson, H., "A decision support framework for automated screening of diabetic retinopathy", International journal of biomedical imaging, vol. 2006(1), Art. No. 045806, 2006, DOI:/10.1155/IJBI/2006/45806

[16] Ainapur, S. S., "Automated diabetic retinopathy detection using a multi-step framework with stacked ensemble-based classification model", Expert Systems with Applications, vol.284, Art. No.127709, 2025, DOI:10.1016/j.eswa.2025.127709

[17] Rafid, A. U. I., Sanjana, S., Munir, M. B., & Sharmin, N., "An early-stage diagnosis of diabetic retinopathy based on ensemble framework", Signal, Image and Video Processing, vol. 18(1), pp. 735-749, 2024, DOI:10.1007/s11760-023-02796-5

[18] Lavanya, K., Thangamani, M., Ganthimathi, M., Priyanka, S., Peter, V. J., & Sapthika, J., "A Multi-Class Deep Neural Network Framework Driven Automated Classification of Diabetic Retinopathy Using Retinal Fundus Images", Genetics and Molecular Research, vol. 25(1), 2026, DOI:10.4238/4j99yb84

[19] Rakhlin, A., "Diabetic retinopathy detection through integration of deep learning classification framework", BioRxiv, Art. No. 225508, 2017, DOI:10.1101/225508

[20] Murugappan, M., Prakash, N. B., Jeya, R., Mohanarathinam, A., Hemalakshmi, G. R., & Mahmud, M, "A novel few-shot classification framework for diabetic retinopathy detection and grading", Measurement, vol.200, Art. No.111485, 2022, DOI: 10.1016/j.measurement.2022.111485

[21] Costa, P., Galdran, A., Smailagic, A., & Campilho, A. "A weakly-supervised framework for interpretable diabetic retinopathy detection on retinal images", IEEE access, vol. 6, pp.18747-18758, 2018, DOI:10.1109/ACCESS.2018.2816003

[22] Alshahrani, M., Al-Jabbar, M., Senan, E. M., Ahmed, I. A., & Saif, J. A. M., "Hybrid methods for fundus image analysis for diagnosis of diabetic retinopathy development stages based on fusion features", Diagnostics, vol. 13(17), pp. 2783. DOI:10.3390/diagnostics13172783

[23] Suganyadevi S., Seethalakshmi V. "FACNN: fuzzy-based adaptive convolution neural network for classifying COVID-19 in noisy CXR images", Med

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Biol Eng Comput, vol. 62, pp. 2893–2909, 2024, DOI:10.1007/s11517-024-03107-x

[24] G. Pradeep, S. Bala, N. P. Satheesh, M. Mahalakshmi, K. Balasamy and S. Suganyadevi, "An Effective Framework for Detecting Epileptic Seizures using CNN and Encrypted EEG Signals", International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, pp. 611-617, 2023, DOI:10.1109/ICSCSS57650.2023.10169196

[25] S. Suganyadevi, K. Renukadevi, K. Balasamy and P. Jeevitha, "Diabetic Retinopathy Detection Using Deep Learning Methods," First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichy, India, pp. 1-6, 2022, DOI:10.1109/ICEEICT53079.2022.9768544.

[26] L. Sowmiya, A. S. Rajasekaran, S. Suganyadevi, S. Sureshkumar, G. Subramaniam and R. Jaazieliah, "A Secure Authenticated Message Transfer in Healthcare Application," 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, pp. 1-6, 2023, DOI:10.1109/ICICACS57338.2023.10099494.

[27] Usha, S., Bala, S., Saranya, M.D. et al. "Pixelated disparity network for hepatocellular carcinoma recognition from ultrasound images", Evolving Systems vol.16, no.113, 2025, DOI:10.1007/s12530-025-09737-2.

[28] Dayana AM, Emmanuel WRS, "An enhanced swarm optimization-based deep neural network for diabetic retinopathy classification in fundus images", MultimedTools Appl, vol. 81, pp.20611–20642, 2022, DOI:10.1007/s11042-022-12492-0

[29] Luo X, Xu Q, Wang Z, Huang C, Liu C, Jin X, Zhang J. "A Lesion-Fusion Neural Network for Multi-View Diabetic Retinopathy Grading", IEEE J Biomed Health Inform, Vol. 29(5), pp. 3184-3193, 2025, DOI:10.1109/JBHI.2024.3384251.

[30] S. Suganyadevi, V. Pavatharani, K. Koushick, R. Manimaran, A. Mohammed Sameer and K. Murugan, "Multisalient Object Detection and Recognition using Deep Learning Techniques," International Conference on Applied Intelligence and Sustainable Computing (ICAISC), Dharwad, India, pp. 1-7, 2023, DOI:10.1109/ICAISC58445.2023.10200299.

[31] Correra, Simona, Valeria Sorgente, Mario Cesarelli, Fabio Martinelli, Antonella Santone, and Francesco Mercaldo, "A Novel Convolutional Neural Network for Explainable Diabetic Retinopathy Detection and Grade Identification" Sensors, vol. 26, no. 8, pp.2510, 2026, DOI:10.3390/s26082510.

[32] S. Suganyadevi, K. R. Sasiprabha, V. Vethavarna and A. S. Rajasekaran, "Multi-Scale Guided Pyramid Transformer with Uncertainty-Aware Deep Refinement for High-Precision Lung Segmentation and Multi-Class Disease Classification," Seventh International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kalyani, India, pp. 675-680, 2025, DOI:10.1109/ICRCICN68210.2025.11364701.

[33] Rahaman, M., Muni, M.A., Tasnim, S. et al. "Explainable AI for diabetic retinopathy detection using vision transformers", Sci Rep vol.16, Art. No. 27354, 2026, DOI:10.1038/s41598-026-56569-4.

[34] Suganyadevi, S., Seethalakshmi, V. & Balasamy, K. "A review on deep learning in medical image analysis. Int J Multimed Info Retr, vol. 11, pp.19–38, 2022, DOI:10.1007/s13735-021-00218-1

[35] Alanazi, Saad, and Rayan Alanazi. "Enhancing diabetic retinopathy detection through federated convolutional neural networks: Exploring different stages of progression." Alexandria Engineering Journal, vol. 120 pp.215-228, 2025, DOI:10.1016/j.aej.2025.02.026.

[36] Suganyadevi, S., Seethalakshmi, V. CVD-HNet: Classifying Pneumonia and COVID-19 in Chest X-ray Images Using Deep Network. Wireless Pers Commun 126, 3279–3303 (2022). https://doi.org/10.1007/s11277-022-09864-y

[37] Niu Y, Gu L, Zhao Y, Lu F. "Explainable Diabetic Retinopathy Detection and Retinal Image Generation" IEEE J Biomed Health Inform, vol. 26(1), pp. 44-55. https://doi.org/10.1109/JBHI.2021.3110593.

[38] Yurdakul, M., Uyar, K., Taşdemír, Ş. & Atabaş, İ. Ropgcvit, "A novel explainable vision transformer for retinopathy of prematurity diagnosis", IEEE Access, vol. 13, pp. 77064–77079, 2025, DOI:10.1109/ACCESS.2025.3564213.

[39] Lalithadevi, B., Krishnaveni, S. "Diabetic retinopathy detection and severity classification using optimized deep learning with explainable AI technique" Multimed Tools Appl vol. 83, pp. 89949–90013, 2024, DOI:10.1007/s11042-024-18863-z

[40] Tashkandi, Araek. "Eye Care: Predicting Eye Diseases Using Deep Learning Based on Retinal Images", Computation, vol. 13, no. 4: pp,91, 2025, DOI:10.3390/computation13040091.

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Author Contribution



Dr. K. Venu is currently serving as an Assistant Professor (Senior Grade) in the Department of Computer Science and Engineering at Kongu Engineering College, Tamil Nadu, India. She completed her Ph.D. in Deep Learning at Anna University, where her research focused on advanced learning models and intelligent computing techniques. With over 9 years of teaching experience, she has actively contributed to both academic instruction and research mentoring. She has published 7 research articles in reputed international journals and presented 11 papers at international and national conferences. In addition, she has authored 3 book chapters with well-known publishers. Her research interests include Machine Learning, Deep Learning, and Nature-Inspired Computing, with applications in intelligent and data-driven systems.



Dr. Thanga Parvathi B is a distinguished Professor at Bannari Amman Institute of Technology, Sathyamangalam, with more than seven years of dedicated experience in engineering education. She holds a Ph.D. in Information and Communication Engineering from Anna University, Chennai, demonstrating her strong research orientation and academic excellence. She completed her Bachelor's and Master's degrees in Computer Science from institutions affiliated with Manonmaniam Sundaranar University, Tirunelveli. Her teaching and research interests focus on advanced computing technologies, information systems, and innovative pedagogical practices. Dr. Thanga Parvathi is committed to mentoring students, promoting academic integrity, and contributing actively to institutional and professional development.



Dr. Kunchanapalli Rama Krishna is a Professor in the CSIT Department at KL University with over 33 years of academic and research experience in Computer Science & Engineering. He holds a Ph.D. in CSE and has served at reputed institutions including Galgotias University, C-DAC, and IIMT College of Engineering. As an AICTE Margdarshak and NAAC e-Assessor, he has guided institutions toward NBA accreditation and academic excellence. He has published over 30 research papers in reputed national and international journals and conferences. Dr. Rama Krishna has also led funded projects and national conferences on cryptography and modernizing education. He authored the book Cloud Computing Fundamentals, showcasing his commitment to knowledge dissemination. He actively participates in FDPs and holds certifications from IITs, IIMs, and NPTEL/SWAYAM. His guidance has shaped numerous UG, PG, and Ph.D. scholars, contributing to quality education and innovation.



Dr. S. Yuvaraj is currently working as an Assistant Professor in the Department of Computer Science and Engineering at Sri Eshwar College of Engineering, Coimbatore, India. He holds a Ph.D. in Computer Science and Engineering from Anna University, Chennai, and has accumulated over 12 years of academic experience in engineering education. His research interests lie in the fields of Cognitive Computing, Machine Learning, Deep Learning, and Artificial Intelligence. He has actively participated in various faculty development programs, seminars, workshops, and both national and international conferences. Dr. Yuvaraj is a professional member of IEEE and is known for his consistent contributions to the academic and research communities. He has published numerous research articles in reputed peer-reviewed international journals, demonstrating a strong commitment to innovation and excellence in computer science research.

Manuscript Received 8 August 2026; Accepted: 14 October 2026; Published online: 4 November 2026
DOI: https://doi.org/10.35882/jeeemi.v9i1.1998
Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).