

DLSTM-121 with Principal Component Analysis for Diabetic Retinopathy Classification Using Segmented Retinal Blood Vessel Images

Muhammad Arhami^{1,5}, Marwan Ramli², Melinda³, Anita Desiani⁴, Musta'inul Abdi⁵ and Husaini⁶

¹ Doctoral Program, School of Engineering, Postgraduate Program, Universitas Syiah Kuala, Aceh, Indonesia

² Department of Mathematics, Universitas Syiah Kuala, Banda Aceh, Indonesia

³ Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Banda Aceh, Indonesia

⁴ Department of Mathematics, Universitas Sriwijaya, Palembang, Indonesia

⁵ Department of Informatics Engineering, Politeknik Negeri Lhokseumawe Aceh, Indonesia

⁶ Department of Computer and Network Engineering Technology, Politeknik Negeri Lhokseumawe Aceh, Indonesia

Corresponding author: Muhammad Arhami (e-mail: Muhammad.arhami@pnl.ac.id), **Author(s) Email:** Marwan Ramli (e-mail: marwan.math@unsyiah.ac.id), Melinda (e-mail: melinda@usk.ac.id), Anita Desiani (e-mail: anita_desiani@unsri.ac.id), Husaini (e-mail: husaini@pnl.ac.id), Musta'inul Abdi (e-mail: mustainul.abdi@pnl.ac.id)

Abstract Diabetic retinopathy classification using fundus images increasingly relies on deep learning, particularly Convolutional Neural Networks (CNNs). However, the use of high-dimensional retinal images in the processing pipeline may increase computational requirements. This study proposes a DLSTM-121-based approach for diabetic retinopathy classification by applying Principal Component Analysis (PCA) to augmented retinal images before U-Net-based segmentation. The main methodological contribution is the placement of PCA before the segmentation stage, allowing dimensionality reduction to be performed before the images are processed by the subsequent segmentation and classification models. DLSTM-121 combines DenseNet-121 and Long Short-Term Memory (LSTM) to extract spatial features and model dependencies among the extracted features. The proposed pipeline consists of data augmentation, PCA, U-Net-based retinal image segmentation, and DLSTM-121 classification. Four PCA configurations using 25, 50, 75, and 100 principal components were evaluated. The results showed that 75 principal components retained approximately 99% of the image information and achieved the best classification performance, with 96.92% accuracy, 100% sensitivity, 86.67% specificity, 98.18% F1-score, and 99% AUC. This configuration required 1,435.59 s of training time and 261.76 MB of peak GPU memory, compared with 1,456.18 s and 232.41 MB for the configuration without PCA. The inference times were 177.22 and 173.20 ms/image, respectively, while both configurations required 12.713 GFLOPs. These results demonstrate that PCA can reduce the dimensionality of the retinal image representation while maintaining high classification performance. The study used 129 retinal images from a single dataset, which limits the generalizability of the findings. Future research should evaluate the proposed approach using larger and more diverse retinal datasets.

Keywords: Diabetic Retinopathy; Deep Learning; Health Risk; Innovation, PCA

1. Introduction

Diabetic retinopathy (DR) is a complication of diabetes that causes damage to the retinal blood vessels and is one of the major causes of visual impairment worldwide. According to the World Health Organization, 422 million people have diabetes, and approximately 35% develop DR [1][2][3]. Early identification of retinal abnormalities is important to support timely treatment and prevent disease progression. Retinal fundus images provide visual information that can be used for automated DR

detection and classification [4]. With the development of deep learning, convolutional neural networks (CNNs) have been widely applied to retinal image classification because of their ability to automatically learn discriminative image features [5][6][7]. DenseNet-121 is one CNN architecture that has demonstrated effectiveness in medical image classification. Its dense connectivity enables feature reuse and facilitates the extraction of rich spatial representations [8].

Several studies have applied DenseNet-121 to diabetic retinopathy classification. Mohanty et al. [9]

achieved an accuracy of 97.30%, while Das and Majumder [10] reported an accuracy above 92%. However, these studies reported accuracy but did not comprehensively evaluate other metrics, such as sensitivity, specificity, and F1-score. Another study reported 79% accuracy using DenseNet-121 for diabetic retinopathy classification [11]. Although these results demonstrate the capability of DenseNet-121 to extract spatial features, CNN-based architectures do not explicitly model sequential dependencies among feature representations [12]. Long Short-Term Memory (LSTM) provides a mechanism to model sequential dependencies and retain relevant information through its memory structure [13][14][15][16]. LSTM has also been applied to retinal fundus image classification. Rahul [17] reported an accuracy of 76% using LSTM for diabetic retinopathy classification, whereas Palaniswamy and Vellingiri [18] achieved an accuracy of 99%, with sensitivity and specificity of 94%, for four-class classification. However, LSTM is not specifically designed to extract spatial image features as effectively as CNN-based architectures. The complementary characteristics of DenseNet-121 and LSTM therefore provide a basis for combining spatial feature extraction and sequential dependency modeling in a hybrid architecture, referred to in this study as DLSTM-121. Although combining DenseNet-121 and LSTM can provide complementary feature learning, processing retinal images with high-dimensional representations through a hybrid deep learning architecture can increase computational requirements. The computational burden may be reflected in training time, memory consumption, and other computational requirements. This issue becomes particularly relevant when computational resources are limited. Therefore, reducing the dimensionality of the retinal image representation is an important consideration for controlling the computational burden while retaining information relevant to classification.

Principal Component Analysis (PCA) is a dimensionality reduction technique that transforms high-dimensional data into a smaller number of principal components while retaining the dominant variance [19][20]. Previous studies have incorporated PCA into diabetic retinopathy classification using different approaches, including ResNet [21], VGG-Net [22], deep neural networks [23], and SIFT [24]. Saeed et al. [21] reported sensitivity and specificity above 90% using PCA with ResNet, while Mateen et al. [22] achieved 92.21% accuracy using PCA with VGG-Net. Gadekallu et al. [23] obtained accuracy, sensitivity, and specificity of 90.07%, 87%, and 91%, respectively, using PCA with a deep neural network. Mansour [24] reported accuracy, sensitivity, and specificity of 91.03%, 97%, and 89%, respectively, using PCA with SIFT. These studies demonstrate the application of PCA in retinal image classification; however, they

employ classification frameworks different from the DLSTM-121 approach considered in this study. In addition to dimensionality reduction, the representation of retinal vascular structures is important for diabetic retinopathy analysis. Because DR is associated with retinal vascular abnormalities, retinal blood vessel segmentation can provide a more focused representation of vascular structures in fundus images. By separating vascular structures from other retinal regions, segmented vessel images can provide a targeted representation for subsequent feature extraction and classification. However, integrating PCA-based dimensionality reduction before retinal blood vessel segmentation with a DenseNet-121–LSTM hybrid architecture has not been sufficiently investigated. In particular, the relationship between the number of retained principal components, computational requirements, and subsequent classification performance within this processing pipeline remains to be examined.

Thus, although DenseNet-121–LSTM architectures can jointly model spatial and sequential characteristics, they do not directly address the computational burden associated with processing high-dimensional retinal image representations, particularly under limited computational resources. In addition, the integration of dimensionality reduction before retinal blood vessel segmentation with DLSTM-121 classification has not been sufficiently investigated. Therefore, to address this problem, this study aims to develop a DLSTM-121-based classifier combined with Principal Component Analysis (PCA) for diabetic retinopathy classification using segmented retinal blood vessel images. The contributions of this study are 1) applying data augmentation and PCA before retinal blood vessel segmentation to obtain a more compact retinal image representation, 2) evaluating 25, 50, 75, and 100 principal components to investigate their effects on classification performance and computational requirements, 3) incorporating retinal blood vessel segmentation to provide a focused representation of retinal vascular structures, and 4) developing DLSTM-121 by combining DenseNet-121 and LSTM to integrate spatial feature extraction and sequential dependency modeling for diabetic retinopathy classification. To achieve this, the augmented retinal images were processed using different PCA configurations and subsequently subjected to retinal blood vessel segmentation. The resulting segmented retinal blood vessel images were then classified using DLSTM-121. Classification performance was evaluated using accuracy, sensitivity, specificity, F1-score, and Cohen's kappa, while computational requirements were assessed using the corresponding computational measures.

This study is structured as follows: Section II discusses the dataset, preprocessing procedures, PCA, retinal blood vessel segmentation, DLSTM-121 architecture, and experimental procedures. Section III presents the results of PCA dimensionality, computational requirements, and diabetic retinopathy classification performance. Section IV discusses the results, compares the proposed approach with previous studies, and presents the limitations of the study. Section V concludes the study by summarizing the main findings and discussing future work.

II. Method

A. STARE Dataset

The STARE (Structured Analysis of the Retina) dataset was developed as part of a retinal image analysis project initiated by Michael Goldbaum, M.D., at the University of California, San Diego, and is publicly available through the STARE project website. The dataset contains retinal fundus images with a resolution of 700×605 pixels in PPM format [25][26]. The STARE dataset contains 400 retinal fundus images with various diagnostic conditions, including Normal, background diabetic retinopathy (BDR), proliferative diabetic retinopathy (PDR), hypertensive retinopathy (HR), branch retinal vein occlusion, central retinal vein occlusion, arteriosclerotic retinopathy, macroaneurysm, and other retinal diseases. In this study, 129 original retinal fundus images were selected based on the diagnosis-code annotations provided in the STARE dataset. The selected images consisted of 36 Normal images, 70 BDR images, and 23 PDR images. For the diabetic retinopathy (DR) classification task, BDR and PDR were combined into a single DR class, resulting in 93 DR images, while the remaining 36 images were categorized as Normal. In addition, 20 images with manually annotated vessel masks were used for the vessel segmentation task. Fig. 1 shows retinal images from the STARE dataset with the segmented vessels of a retinal image.

B. Affine Scaling

In the preprocessing stage, retinal images from the STARE dataset, originally with a resolution of 700×605 pixels, were resized to meet the model input requirement of 336×336 pixels. This process was initiated by determining the scaling factor using an affine scaling approach. The scaling factor was computed by comparing the ratio between the target input dimensions and the original image dimensions along each axis using Eq. (1) [27]:

$$s = \min \left(\frac{H_{target}}{H_{input}}, \frac{W_{target}}{W_{input}} \right) \quad (1)$$

where H_{input} and W_{input} denote the height and width of the original image, while H_{target} and W_{target} represent the height and width of the target input. The variable s

corresponds to the scaling factor applied during resizing. The smaller ratio was selected to ensure that the entire image fits within the target dimensions without cropping. The obtained scaling factor is used to resize the original image using Eq. (2) and Eq. (3) [27]:

$$H_{new} = H_{input} * s \quad (2)$$

$$W_{new} = W_{input} * s \quad (3)$$

H_{new} and W_{new} denote the height and width of the image after the resizing process, respectively. The results of this calculation indicate that the image height matches the target size of 336 pixels, while the image width remains 290 pixels. Since the scaling process preserves the aspect ratio, the shape of the objects in the image does not undergo any alteration or distortion, such as stretching or compression. To match the required input size, padding was applied along the width dimension to compensate for the missing pixels until it reached 336 pixels. The padding value was determined based on the absolute difference between the target width and the resized width using Eq. (4) [27]:

$$pad = \frac{|H_{new} - W_{new}|}{2} \quad (4)$$

where pad represents the size of the padding to be applied. The size is symmetrically divided into 23 pixels

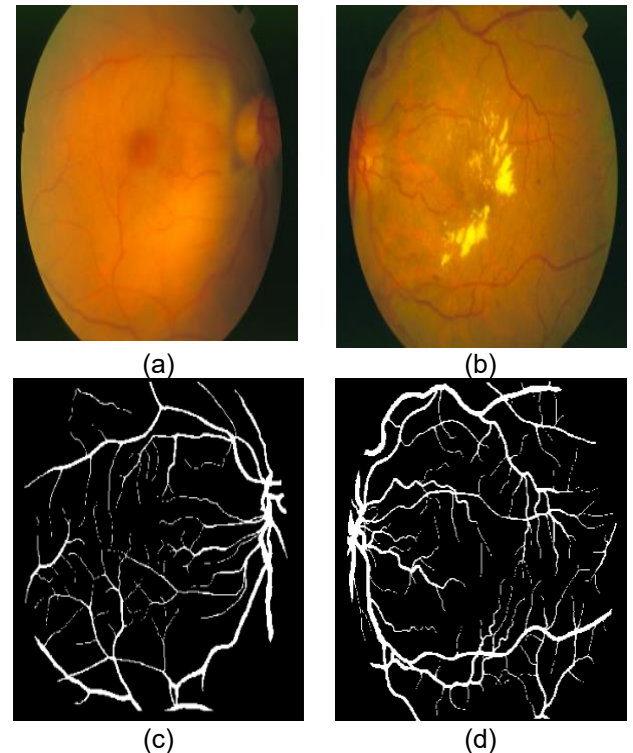


Fig. 1. Representative STARE retinal images and corresponding vessel segmentation masks: (a) normal retinal image, (b) abnormal retinal image, (c) segmented vessels from a normal image, and (d) segmented vessels from an abnormal image

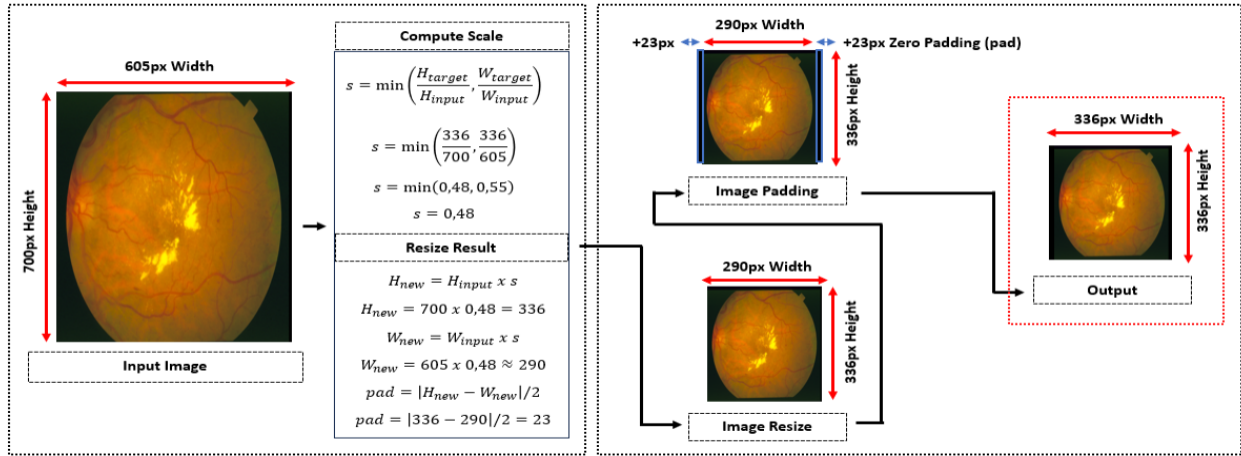


Fig. 2. Affine scaling and zero-padding procedure used to resize retinal fundus images to 336 × 336 pixels while preserving the original aspect ratio.

on the left side and 23 pixels on the right side. Padding is performed using the zero-padding method, which involves adding zero-valued pixels to the additional regions. Consequently, the final image size becomes 336 × 336 pixels without altering the essential information, while preserving the original aspect ratio of the image used as the input. All affine scaling processes are shown in Fig. 2.

C. Enhancement of Image Quality

Image quality enhancement is needed to increase the accuracy of segmentation of the DR classification results on retinal images. This study applies the green channel of the image and contrast-limited adaptive histogram equalization (CLAHE) to enhance image contrast. In Fig. 3, the resulting images in each image channel have different contrasts. The red channel produces images with the brightest contrast. The blue channel has the lowest contrast, resulting in darker images. The green channel produces images with the

best contrast. In this study, images from the green channel were selected to be contrast-enhanced using CLAHE so that the resulting images are expected to have good quality with appropriate contrast intensity. CLAHE operates on local regions of an image, commonly referred to as tiles, by limiting the histogram within each region to control excessive contrast enhancement. The clip limit is determined based on the tile area, the number of gray levels, the clip factor, and the maximum allowable slope, as shown in Eq. (5) [28]:

$$E = \frac{M}{N} \left(1 + \frac{\alpha}{100} (S_{max} - 1)\right) \quad (5)$$

where E represents the clip limit, M represents the area of the tile, N represents the number of gray levels, α represents the clip factor, and S_{max} represents the maximum allowable slope. The histogram values in each tile are clipped according to the calculated clip limit, and the adjacent tiles are subsequently combined using bilinear interpolation.

D. Dataset Splitting

After image preprocessing, including affine scaling and CLAHE enhancement, the 129 original retinal fundus images were divided into training, testing, and validation subsets using an 80:10:10 ratio. Accordingly, 103 images were assigned to the training set, 13 images to the testing set, and 13 images to the validation set. The training set consisted of 29 Normal, 57 BDR, and 17 PDR images. The testing set consisted of 3 Normal, 7 BDR, and 3 PDR images, while the validation set consisted of 4 Normal, 6 BDR, and 3 PDR images. For the DR classification task, BDR and PDR were considered as the DR class.

E. Data Augmentation

Implementation of data augmentation can increase the amount of data available for the training model without requiring new data. Data augmentation can increase

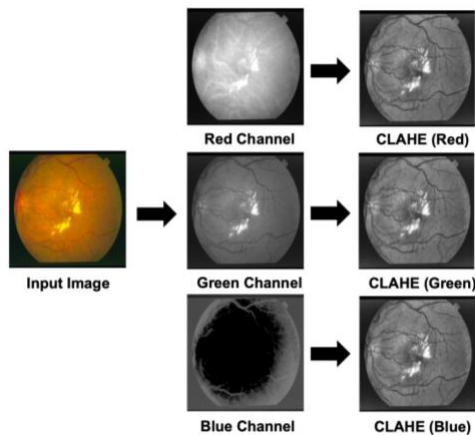


Fig. 3. CLAHE-based contrast enhancement of retinal images in the red, green, and blue color channels.

the number of training samples without losing semantic information and allows image data to be free from bias.

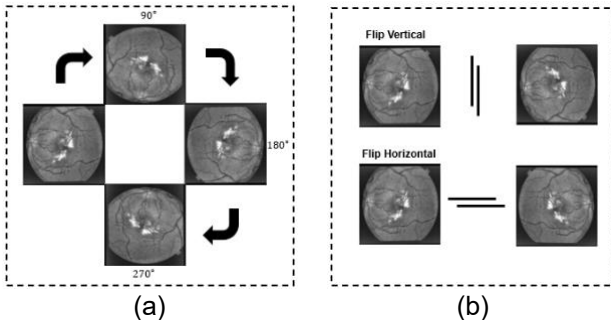


Fig. 4. Examples of training-data augmentation: (a) retinal image rotation at different angles and (b) vertical and horizontal image flipping.

Generally, bias occurs because of the similarity between the image data [8]. Bias can be overcome by making basic modifications to the image, such as random rotation and flipping [5]. Fig. 4 presents an illustration of the data augmentation process on retinal images, and the image is modified by rotating and flipping. Image rotation is performed with random clockwise angles. In Fig. 4a, the image rotation results are given with angular displacements of 90°, 180°, and 270°. Image flipping was performed vertically and horizontally in Fig. 4b. Data augmentation was subsequently applied only to the training set, while the validation and testing sets retained the original images without augmentation. This procedure was implemented after the dataset was divided into training, validation, and testing subsets to prevent data leakage and ensure that no augmented version of an image in the training set was introduced into the validation or testing sets. By restricting augmentation to the training data, the validation and testing sets remained independent and could provide an unbiased evaluation of the model's generalization performance.

F. Dimension Reduction using PCA

Principal component analysis (PCA) is a multivariate analysis method used to reduce data dimensions, support big data visualization, image compression, data mining, machine learning, and monitoring processes [29]. In the PCA process, input image data are converted into a high-dimensional feature space [30]. Based on mathematical calculations, the PCA process is conducted by reducing the number of n -dimensional feature spaces to k -dimensional ($k < n$). k -dimensional feature space is a new orthogonal feature called a principal component (PC) from the reconstruction of the n -dimensional feature space [31]. In this study, PCA was applied only to the training data after dataset splitting and data augmentation to reduce the dimensionality of the retinal images before the U-Net

segmentation process. PCA can be performed using the following steps:

1. Data Standardization

The data-standardization process can avoid the influence of different feature-space dimensions [32]. shows the Data standardization Eq. (6) [33].

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (6)$$

where x_{ij} is the input image data entry, $\bar{x}_j$ is the average value of the j -feature space, σ_j is the standard deviation of the j -feature space, and z_{ij} is new entry data. Matrix entries obtained from the standardization process are entered into the data standardization result matrix (Z).

2. Covariance Matrix

The covariance matrix is a symmetrical square matrix of order $n \times n$. Eq. (7) shows the covariance matrix [34]:

$$C = \begin{bmatrix} \sigma_{11}^2 & Cov(Z_1, Z_2) & \cdots & Cov(Z_1, Z_j) \\ Cov(Z_2, Z_1) & \sigma_{22}^2 & \cdots & Cov(Z_2, Z_j) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(Z_i, Z_1) & Cov(Z_i, Z_2) & \cdots & \sigma_{ij}^2 \end{bmatrix} \quad (7)$$

3. Eigen Values and Eigenvectors

Eigenvalues and eigenvectors are important factors in determining the number of principal components, which can be calculated using Eq. (8) [33].

$$Cv = \lambda v \quad (8)$$

where λ is the eigenvalue, and v is the eigenvector. The obtained eigenvalues are sorted from the largest eigenvalues [32].

4. Principal Component

The number of principal components (PC) retained in the retinal image is determined by calculating the PC matrix and cumulative proportion using Eq. (9) and Eq. (10) [32].

$$PC = Z^T V \quad (9)$$

$$P_k(PC_i) = \frac{\lambda_i}{\sum_{i=1}^m \lambda_i} \quad (10)$$

where $i = 1, 2, \dots, m$, λ_i is the eigenvalue of the i , $P_k(PC_i)$ is the cumulative proportion of the i , Z^T is the data standardization transpose matrix, V is the eigenvector matrix, and PC is the principal component matrix. For the training data, PCA was performed by first standardizing the retinal image data using Eq. (6). The standardized training data were then used to construct the covariance matrix, as shown Eq. (7). Based on the covariance matrix, the eigenvalues and eigenvectors were calculated using Eq. (8), and the number of principal components was determined using Eq. (9) and Eq. (10). Finally, the dimensionality of the training images was reduced based on the retained

principal components. The resulting reduced images were subsequently used as input to the U-Net segmentation model. Fig. 5 shows a series of image dimension reduction processes using PCA.

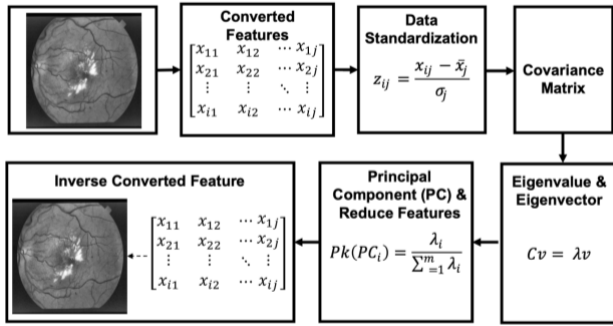


Fig. 5. PCA-based dimensionality-reduction workflow for retinal images, including standardization, covariance computation, eigenvalue–eigenvector decomposition, principal-component selection, and image reconstruction.

G. Image Segmentation using U-Net

In the segmentation process, the image is divided into several regions, such that the pixels contained in one region have the same characteristics. In retinal images, the image is divided into two regions: retinal blood vessels and the background [35]. The retinal image segmentation process is conducted by applying a modified CNN network that resembles a U-shape, namely, U-Net. U-Net can perform end-to-end image segmentation through encoder and decoder lines [36]. Fig. 6 presents the proposed U-Net architecture for image segmentation. The U-Net produces a binary segmented image containing retinal blood vessels and background. The segmented vessel image is used as

the input for DR classification. In this study, the retinal blood vessel representation refers to the visual information contained in the segmented vessel structures rather than manually extracted features such as vessel diameter, tortuosity, or branching patterns. Abnormalities in retinal blood vessel structures can provide information related to the presence of diabetic retinopathy. Therefore, the proposed method focuses on the information contained in retinal blood vessels for DR classification. The main operations in the proposed U-Net architecture consist of convolution, batch normalization (BN), ReLU activation, and pooling. The convolution operation extracts spatial features from the input feature map and is defined as Eq. (11) [37]:

$$Y_{ij} = \sum_{m=1}^M \sum_{n=1}^N P_{(i+m-1)(j+n-1)} K_{mn} + b \quad (11)$$

where P_{ij} is the input matrix (such as a feature map), K is the filter matrix or convolution kernel, b is the bias, and Y_{ij} represents the generated feature maps at spatial location ij . BN is applied to normalize the feature maps after convolution and improve the stability of the training process. The BN operation is formulated as Eq. (12), Eq. (13), and Eq. (14) [38]:

$$\mu_{ij} = \frac{1}{|B_{ij}|} \sum_{Y_{ij} \in B_{ij}} Y_{ij} \quad (12)$$

$$\sigma_{ij}^2 = \frac{1}{|B_{ij}|} \sum_{Y_{ij} \in B_{ij}} (Y_{ij} - \mu_{ij})^2 \quad (13)$$

$$D_{ij} = \gamma_{ij} \frac{Y_{ij} - \mu_{ij}}{\sqrt{\sigma_{ij}^2 + \epsilon}} + \beta_{ij} \quad (14)$$

where B_{ij} denotes the set of input activations considered for batch normalization at layer ij , Y_{ij}

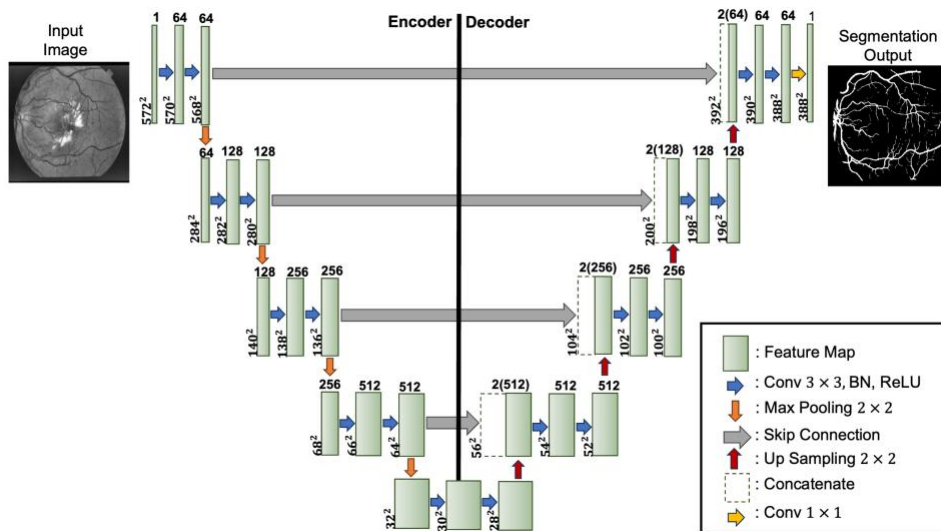


Fig. 6. Proposed U-Net architecture for retinal blood-vessel segmentation, showing the encoder, decoder, skip connections, and segmentation output.

represents an input activation, and $|B_{ij}|$ denotes the number of activations used to calculate the batch statistics. The variables μ_{ij} and σ_{ij}^2 denote the mean and variance of the input activations, respectively. The parameter ε is a small constant introduced to prevent numerical instability, while γ_{ij} and β_{ij} are learnable scale and shift parameters, respectively. Following

where G_i represents the ground-truth value of pixel i , P_i denotes the predicted segmentation value at pixel i , and N is the total number of pixels. Binary Cross-Entropy (BCE) also measures the pixel-wise difference between the predicted segmentation and the ground truth. The BCE loss is defined as Eq. (18) [39]:

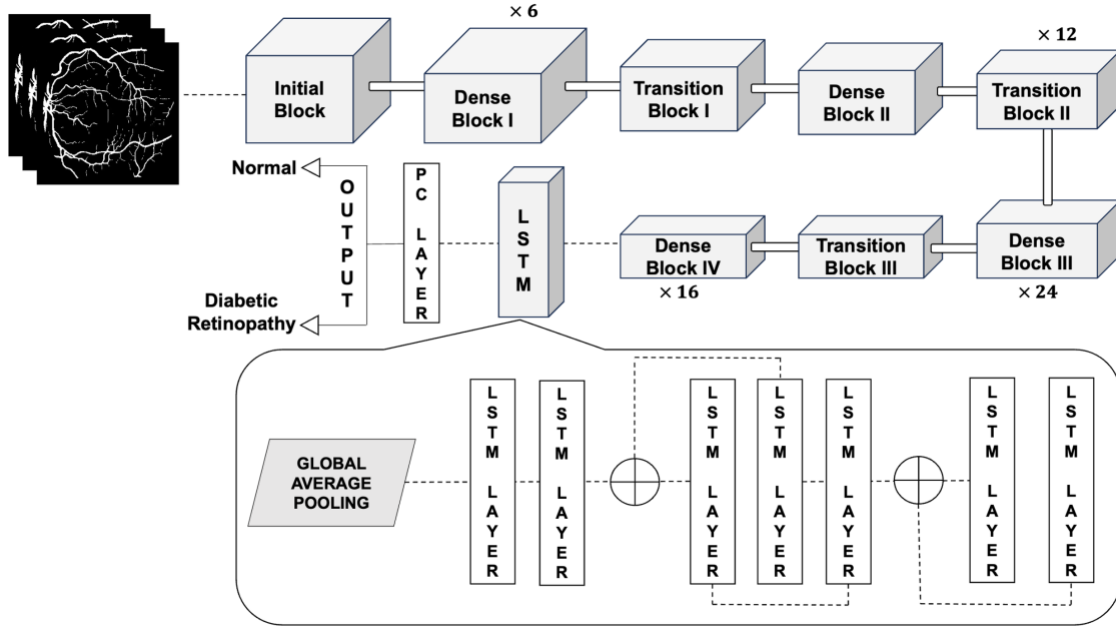


Fig. 7. Proposed DLSTM-121 architecture for diabetic retinopathy classification using DenseNet-121 feature extraction and LSTM-based dependency modeling.

batch normalization, the feature maps pass through the ReLU activation function to introduce non-linearity into the network. The ReLU activation function is defined as Eq. (15) [37]:

$$R_{ij} = \text{ReLU}(D_{ij}) = \max(0, D_{ij}) \quad (15)$$

where Y_{ij} denotes the BN output at spatial position i, j , and R_{ij} represents the activated feature map. The activated feature maps are subsequently processed using max pooling to reduce their spatial dimensions while retaining the most prominent features. The max pooling operation is defined as Eq. (16) [37]:

$$F_{ij} = \max_{(m,n) \in W} R_{(i+m, j+n)} \quad (16)$$

where R denotes the feature map obtained after ReLU activation, F_{ij} represents the max-pooled feature at spatial position ij , and W denotes the pooling window. The segmentation performance is evaluated using Dice Loss, which measures the overlap between the predicted retinal vessel segmentation and the ground-truth mask. The Dice Loss is formulated as Eq. (17) [39]:

$$\mathcal{L}_{DICE} = 1 - \frac{2 \sum_{i=1}^N G_i P_i}{\sum_{i=1}^N G_i + \sum_{i=1}^N P_i} \quad (17)$$

$$\mathcal{L}_{BCE} = -\frac{1}{N} \sum_{i=1}^N [G_i \ln(P_i) + (1 - G_i) \ln(1 - P_i)] \quad (18)$$

where G_i represents the ground-truth label and P_i denotes the predicted probability for pixel i .

H. Classification with DLSTM-121

At the classification stage, the segmented retinal blood vessel image is the input to DenseNet-121, which extracts deep spatial feature maps; the resulting feature representation is then transformed into a feature vector and processed by the LSTM component. Therefore, the features modeled by LSTM are deep features extracted from the segmented vessel representation, rather than manually extracted vascular descriptors. Preprocessing was applied to the retinal images to maximize the performance of the proposed method for classification. A series of preprocessing stages has been conducted, including image quality enhancement, data augmentation, dimension reduction using PCA, and image segmentation using U-Net. The proposed DR classification method was designed using CNN and RNN architectures. For the CNN, the proposed method used DenseNet-121. For the RNN, the proposed method used an LSTM architecture. The data training

process began with training the DenseNet-121 architecture, which had three main blocks. The dense block had four parts; each part was repeated 6, 12, 24, and 16 times, and each dense block was connected by a transition block. After training the DenseNet-121 architecture, the training process continued on the LSTM architecture. The feature map generated by DenseNet-121 was reorganized into an ordered spatial sequence before being provided as input to the LSTM architecture. The output of the LSTM architecture can be transformed into a 1×1 matrix using global average pooling. The final stage of classification was performed in the fully connected layer by applying the sigmoid activation function to determine the state of the retinal image. DR classification is performed by determining the state of the retinal image based on two classes,

Fig. 8 presents a series of DenseNet-121 operations performed on each block for DR classification. In Fig. 8, the DenseNet-121 architecture begins with the initial block, which first performs the zero-padding technique. The zero-padding technique can equalize image size by adding zero entries around the input matrix [40]. After equalizing the image size, a two-dimensional convolution operation was performed with a 7×7 input. The results of Conv2D can be normalized using batch normalization (BN) data, followed by the application of the ReLU activation function to eliminate negative values in the input matrix entries [40]. The pooling operation was performed to maintain the highest matrix entry via MaxPooling2D with a 3×3 input. In a dense block, two-dimensional convolution operations are performed with a 3×3 input

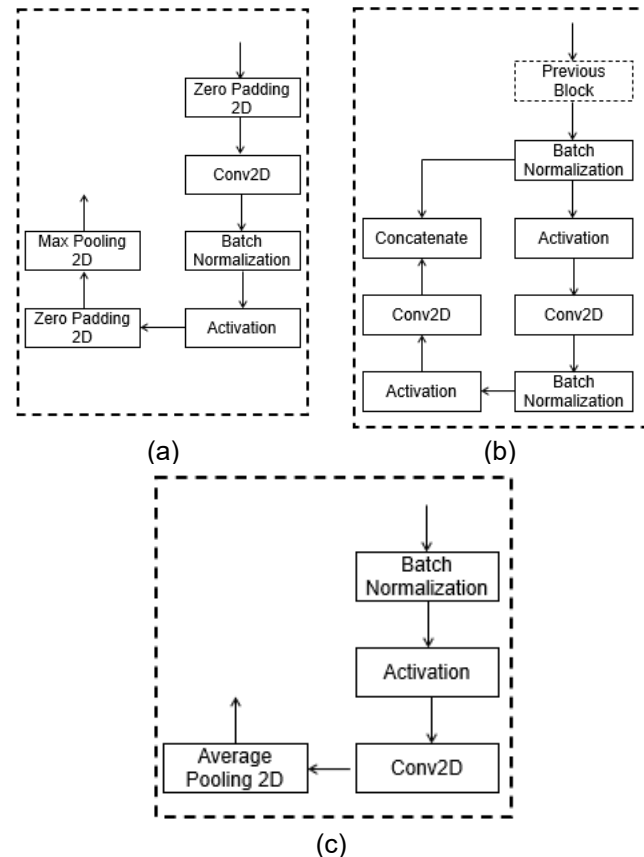


Fig. 8. Main operations in the DenseNet-121 architecture: (a) initial block, (b) dense block, and (c) transition block.

namely, retinal images that are in the normal class and retinal images that are classified as having DR disorders. Fig. 7 shows all steps of the proposed method for DR classification. Fig. 7 has three main blocks in DenseNet-121: the initial, dense, and transition blocks.

Table 1. Configuration for Converting DenseNet-121 Feature Maps into LSTM Input Sequences.

Parameter	Configuration
DenseNet-121 feature map	10 × 10 × 1024
Spatial locations	10 × 10 = 100
Sequence representation	100 × 1024
Sequence length	100
Feature dimension	1024
Spatial ordering	Row-major: left-to-right, top-to-bottom
LSTM layers	1
LSTM units	128
Dropout	0.5
Sequence interpretation	Ordered spatial regions, not temporal observations

to extract image features and a 1×1 input to change the filter size [41]. BN, ReLU, and Conv2D were performed sequentially and were repeated twice. Merging of image features uses the concatenation of a series of operations on dense blocks and can be passed into the transition block. Inside the transition block, a two-dimensional convolution operation was performed with only a 1×1 input. The same operation on the dense block is performed on the transition block, but only once. A pooling operation retrieved the average matrix entry via AveragePooling2D with a 2×2 input at the end of the transition block.

After performing a series of operations on the dense blocks and transition blocks, the feature map generated by DenseNet-121 was reorganized into an ordered spatial sequence before being provided as input to the LSTM architecture. The resulting feature map is 10 × 10 × 1024, where 10 × 10 represents the spatial locations, and 1024 represents the feature dimension at each location. Instead of converting the

feature map into a single 102400-dimensional vector, the 10×10 spatial locations were arranged into 100 sequence elements, resulting in a representation of 100×1024 . The spatial locations were ordered using a row-major arrangement, in which the features were read from left to right within each row and then continued from the top row to the bottom row. Therefore, the sequence length provided to the LSTM was 100, with 1024 features for each sequence element. The sequence does not represent temporal observations but an ordered arrangement of spatial regions extracted from the feature map. This arrangement allows the LSTM to process the spatial feature representations sequentially and retain information from previously processed regions through its hidden and cell states. The resulting feature-map conversion and LSTM input configuration are summarized in Table 1.

As shown in Table 1, the $10 \times 10 \times 1024$ feature map is converted into a sequence of 100 spatial regions, with 1024 features for each region. The sequence is arranged using row-major ordering and is processed by one LSTM layer with 128 units and a dropout rate of 0.5. Thus, the table specifies the configuration used to transform the spatial features generated by DenseNet-121 into the LSTM input representation. Based on this configuration, the LSTM receives a sequence with 100 elements, where each element contains 1024 features. The LSTM consists of one layer with 128 units and a dropout rate of 0.5. At each sequence step, the LSTM receives the feature representation of one spatial region and updates its hidden and cell states. This enables information from previously processed spatial regions to be retained when processing subsequent regions. Therefore, the LSTM is used to model dependencies among the spatial feature representations extracted by DenseNet-121 rather than temporal dependencies. The LSTM architecture has three main gates: the forget gate, the input gate, and the output gate. The LSTM layer is given input in the form of a feature vector that has a hidden state vector and a cell state vector. The LSTM unit is implemented in hardware; therefore, the LSTM is made up of four gates, which are added to a computation gate. Each gate has a matrix and vector multiplication, followed by the addition of a bias vector and the application of the sigmoid and tanh functions.

The LSTM computation starts by initializing the weight matrix of the input state and the weight matrix of the hidden state, where each weight matrix has a dimension of and the entries are randomly generated by the computer. The forget gate is calculated using the sigmoid activation function as defined in Eq. (19) [42].

$$f_t = \sigma(W_{xi}x_t + W_{hf}h_0 + b_f) \quad (19)$$

The weight matrix of the input state (W_{xi}) and the weight matrix of the hidden state (W_{hi}) are initialized

with the same matrix. The input gate (i_t) is calculated using the sigmoid function as shown in Eq. (20) [42].

$$i_t = \sigma(W_{xc}x_t + W_{hc}h_0 + b_i) \quad (20)$$

After that, the weight matrix of the input state (W_{xc}) and that of the hidden state (W_{hc}) are initialized to calculate the candidate cell state. The candidate cell state vector ($\tilde{c}_t$) is obtained using the hyperbolic tangent function as formulated in Eq. (21) [42].

$$\tilde{c}_t = \tanh(W_{xc}x_t + W_{hc}h_0 + b_c) \quad (21)$$

The candidate cell state vector ($\tilde{c}_t$) updated by combining the forget gate and the input gate. This process allows the model to retain relevant information as defined in Eq. (22) [42].

$$c_t = f_t \odot c_0 + i_t \odot \tilde{c}_t \quad (22)$$

The weight matrix of the input state (W_{xo}) and the weight matrix of the hidden state (W_{ho}) are initialized to compute the output gate (o_t) where is calculated using the sigmoid activation function as given in Eq. (23) [42].

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_0 + b_o) \quad (23)$$

Finally, the hidden state vector (h_t) is updated based on the output gate and the updated cell state. The hidden state update is expressed using the hyperbolic tangent function as shown in Eq. (24) [42].

$$h_t = o_t \odot \tanh(c_t) \quad (24)$$

The hidden state vector obtained from the LSTM is passed to the fully connected layer to generate the classification score. The output of the fully connected layer is then transformed into a probability using the sigmoid activation function for binary DR classification. The fully connected operation is defined as Eq. (25) [43]:

$$z = W_{fc}h_t + b_{fc} \quad (25)$$

where h_t denotes the hidden state vector generated by the LSTM, W_{fc} represents the weight matrix, b_{fc} denotes the bias vector, and z represents the classification score. The sigmoid function is subsequently applied to obtain the probability of the input retinal image belonging to the DR class using Eq. (26) [44]:

$$\hat{p} = \frac{1}{1 + e^{-z}} \quad (26)$$

where $\hat{p}$ is the predicted probability of the DR class.

I. Evaluation

The performance of the proposed method was evaluated using Accuracy, Precision, Sensitivity, Specificity, Dice coefficient, Intersection over Union (IoU), F1-score, and Area Under the Curve (AUC). For binary classification, the evaluation metrics were calculated based on the numbers of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Accuracy measures the proportion of correctly classified samples among all evaluated samples and is defined as Eq. (27) [45]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (27)$$

Precision measures the proportion of correctly predicted positive samples among all samples predicted as positive using Eq. (28) [45]:

$$Precision = \frac{TP}{TP + FP} \quad (28)$$

Sensitivity measures the ability of the model to correctly identify positive samples using Eq. (29) [46]:

$$Sensitivity = \frac{TP}{TP + FN} \quad (29)$$

Specificity measures the ability of the model to correctly identify negative samples using Eq. (30) [45]:

$$Specificity = \frac{TN}{TN + FP} \quad (30)$$

The Dice coefficient measures the degree of overlap between the predicted segmentation mask and the ground-truth mask and is defined as Eq. (31) [47]:

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (31)$$

IoU measures the ratio between the intersection and union of the predicted and ground-truth regions using Eq. (32) [47]:

$$IoU = \frac{TP}{TP + FP + FN} \quad (32)$$

The F1-score represents the harmonic mean of Precision and Sensitivity and is calculated as in Eq. (33) [45]:

$$F1 = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \quad (33)$$

The AUC represents the area under the Receiver Operating Characteristic (ROC) curve and summarizes the discriminative ability of the classifier across different classification thresholds. The ROC curve is defined by

the relationship between the true positive rate (TPR) and false positive rate (FPR) using Eq. (34) and Eq. (35) [46]:

$$TPR = \frac{TP}{TP + FN} \quad (34)$$

$$FPR = \frac{FP}{FP + TN} \quad (35)$$

The AUC is then calculated as the area under the ROC curve using Eq. (36) [46]:

$$AUC = \int_0^1 TPR(FPR)d(FPR) \quad (36)$$

where TP , TN , FP , and FN denote true positive, true negative, false positive, and false negative, respectively.

Fig. 9 shows the proposed method flowchart in the study. The proposed method consists of image quality enhancement, data augmentation, dimension reduction using PCA, image segmentation using U-Net, and image classification using DLSTM-121 architecture.

III. Result

A. Experimental Setup

1. Segmentation Training Configuration

The retinal vessel segmentation model was trained using the Adam optimizer with a learning rate of 1×10^{-4} and a batch size of 4 for a maximum of 60 epochs. The loss function combined Binary Cross-Entropy and Dice Loss to improve segmentation performance. Model weights were initialized using the Normal initialization method. The segmentation model was trained using 20 manually annotated vessel masks from the STARE dataset, and training augmentation was applied with an augmentation factor of eight on the training segmentation data. Early stopping with a

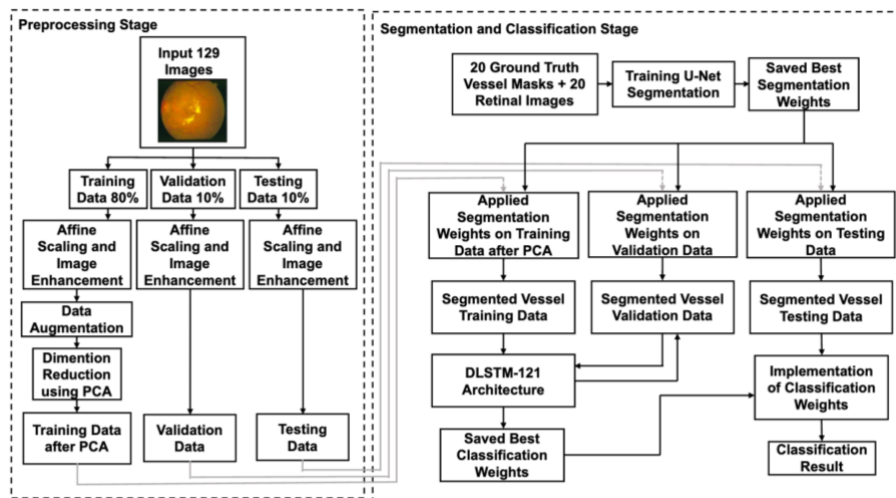


Fig. 9. Overall workflow of the proposed diabetic retinopathy classification framework, including preprocessing, PCA-based dimensionality reduction, U-Net vessel segmentation, and DLSTM-121 classification.

Table 3. Examples of Retinal Images and Vessel Masks Before and After Enhancement and Data Augmentation.

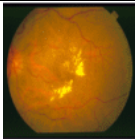
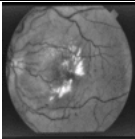
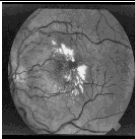
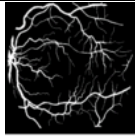
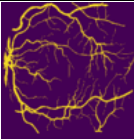
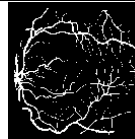
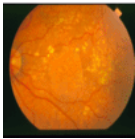
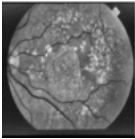
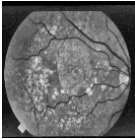
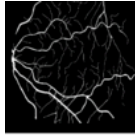
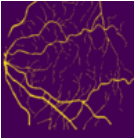
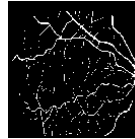
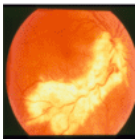
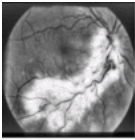
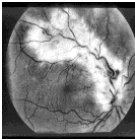
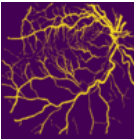
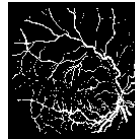
No.	Original Image	Image Enhancement	Image Augmentation	Original Mask	Mask Enhancement	Mask Augmentation
1.						
2.						
3.						

Table 2. Computational Complexity and Training Configuration of the Evaluated Classification Architectures.

PCA	Trainable (M)	Total Params (M)	Inference (ms/Image)	FLOPs (G)	PCA Fit (s)	PCA Size (MB)
PCA25	1.518	7.752	194.70	13.642	0.6158	0.1265
PCA50	1.467	7.694	185.93	13.199	0.5472	0.1128
PCA75	1.405	7.627	177.22	12.712	0.4813	0.1003
PCA100	1.498	7.723	188.45	13.476	0.5669	0.1169
NoPCA	1.624	7895	205.77	14.598	0.7842	0.1548

patience of 8 epochs and the ReduceLROnPlateau learning-rate scheduler were employed to improve training stability. Model checkpointing was based on the minimum validation loss criterion. Performance evaluation was conducted using a 5-fold out-of-fold validation strategy, and the reported metrics included Dice coefficient, Intersection over Union (IoU), sensitivity, specificity, and 95% confidence intervals.

2. Classification Training Configuration

The classification model was trained using the Adam optimizer with an initial learning rate of 1×10^{-3} , which was reduced to 1×10^{-5} during the fine-tuning stage. The model was trained with a batch size of 16 for a maximum of 100 epochs using the Binary Cross-Entropy loss function. DenseNet121 was initialized with ImageNet pre-trained weights and combined with a single LSTM layer consisting of 128 units and a dropout rate of 0.5. The training process used 15,000 augmented training images, resulting in 938 steps per epoch. Early stopping was applied with a patience of 10 epochs by monitoring the validation AUC, while the best model was selected using the maximum validation AUC criterion. The transfer-learning strategy consisted of an initial training phase with the DenseNet121 backbone frozen, followed by fine-tuning of the last 40 layers using

a lower learning rate. No augmentation was applied to the validation and testing datasets.

B. Data Augmentation

This study applied augmentation techniques to increase the amount of image data, and PCA methods were used to reduce image data dimensions. Augmentation was performed on the STARE dataset using image flipping and image rotation between 0 and 360° randomly. An augmentation application increases the amount of data. Augmentation techniques can produce images different from those of the original image. Table 2 presents image augmentation results. The augmentation technique was then applied to the training data, increasing the number of training images to 15,000, while the validation and testing data remained unchanged. These data were subsequently used for retinal vessel segmentation using U-Net and diabetic retinopathy classification using DenseNet121-LSTM.

C. Dimension Reduction and Segmentation

Dimension reduction was performed using Principal Component Analysis (PCA) to obtain a more compact representation of the retinal images while retaining the main information in the data. Five configurations were considered, namely NoPCA, PCA25, PCA50, PCA75, and PCA100. NoPCA represents the original retinal

images without dimensionality reduction, while the PCA configurations represent images reduced to 25, 50, 75, and 100 principal components, respectively. The results showed that increasing the number of principal components increased the amount of information retained in the retinal images. PCA25 retained 94.26% of the variance, while PCA50, PCA75, and PCA100 retained 97.21%, 98.41%, and 99.03%, respectively. However, using more components also increased the representation size and PCA fitting time. The computational complexity and training configurations of the evaluated architectures are summarized in Table 3. The results show that PCA75 retained 98.41% of the variance with a representation size of 0.1003 MB and a fitting time of 0.4813 s. Compared with PCA100, which retained 99.03% of the variance, PCA75 lost only 0.62 percentage points of variance while requiring a smaller representation size and shorter fitting time. PCA75 was selected for the subsequent U-Net segmentation stage.

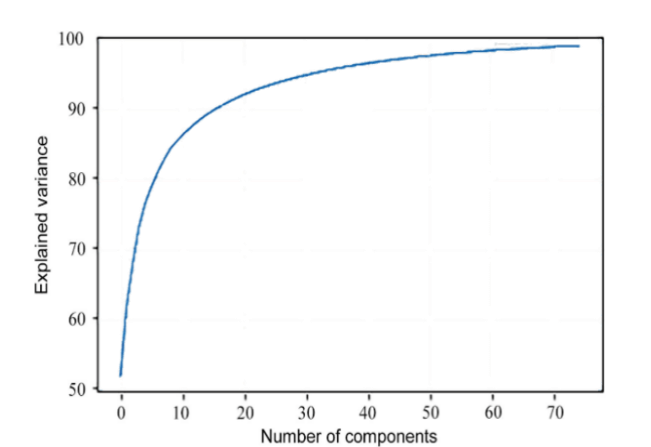


Fig. 10. Cumulative explained variance as a function of the number of retained principal components for PCA-based dimensionality reduction.

Table 4. Segmentation and Classification Performance for Different Numbers of PCA Components

PCA	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC (%)	Dice (%)	IoU (%)
NoPCA	84.62	80	100	88.89	93	62.71	49.42
PCA25	69.23	70	66.67	77.78	70	38.54	24.00
PCA50	76.92	80	66.67	84.21	80	55.11	38.23
PCA75	96.92	100	86.67	98.18	99	69.75	53.72
PCA100	84.62	90	66.67	88.89	93	66	45.84

The segmented retinal images obtained from this stage were then used as input for the DLSTM-121 classification process. The explained-variance pattern is also visible in Fig. 10. The curve rises sharply at the beginning: the first component explains about 52% of the variance, around 75% is reached with five components, and approximately 85-90% is reached within the first 10-15 components. After about 20 components, the increase becomes more gradual, and the curve starts to level off after around 30 components. With 75 components, the explained variance reaches 98.41%, indicating that most of the variation in the original data is already retained. Overall, the results show that PCA can reduce the dimensionality of the retinal image data while preserving most of the information. However, the optimal configuration should not be selected based solely on explained variance. Each PCA configuration was therefore evaluated as input to U-Net. The resulting segmentation performance is shown in Table 4.

Table 4 shows that segmentation performance varies across the PCA configurations. Compared with PCA25, increasing the number of components to PCA50 improves the Dice score and IoU, indicating that more components preserve useful vessel information. PCA75 provides the best overall segmentation performance,

with an accuracy of 96.92%, sensitivity of 100%, Dice of 69.75%, and IoU of 53.72%. Compared with NoPCA, PCA75 improves both accuracy and Dice score. This improvement may reflect PCA retaining sufficient vessel information while removing redundant variations in the image representation. In contrast, PCA25 and PCA50 may discard useful vessel-related information, while the additional components in PCA100 do not provide further segmentation benefits. PCA75 provides the most suitable representation for the U-Net segmentation stage. The training behavior of U-Net is presented in Fig. 11. Training and validation accuracy increase rapidly during the early epochs and then remain relatively stable at around 0.93–0.95. The training accuracy stays slightly above the validation accuracy, but the gap remains relatively small. The training and validation losses generally decrease over the course of training, although some fluctuations can be seen in the validation loss toward the later epochs. Overall, the similar training and validation curves indicate that the model maintains consistent segmentation performance across the two datasets. Precision and recall during U-Net training are shown in Fig. 12. Training precision reaches approximately 0.98, while validation precision remains around 0.96. Training recall is also around 0.98, with validation recall around 0.97. The relatively small

differences between the training and validation curves indicate that the model maintains similar performance when applied to the validation data.

The segmentation results generally follow the main blood vessel patterns visible in the ground truth. The main vessel structures can be identified in the predicted masks and are also visible in the overlay images. These visual results support the quantitative findings in Table 4, particularly the strong performance obtained with the PCA75 configuration. The results from Table 3, Fig. 11, Fig. 12, and Table 4 indicate that PCA75 combined with U-Net provides the most suitable segmentation configuration among those evaluated. PCA75 retains most of the image information while producing the

applied to the retinal images, with 75 principal components selected based on the previous results. The resulting images were processed using U-Net, and the segmented images were then used as input for DLSTM-121 classification. The classification process begins with a 336×336 pixel image. Zero-padding was applied, followed by feature extraction using Conv2D, as shown in Fig. 13. The extracted features were normalized using batch normalization (BN) and ReLU activation, followed by processing through dense blocks and feature concatenation. Fig. 13a shows the first convolution layer results, while Fig. 13b shows the concatenated layer result in the first dense block for DR classification. The resulting features were then used as input to the LSTM

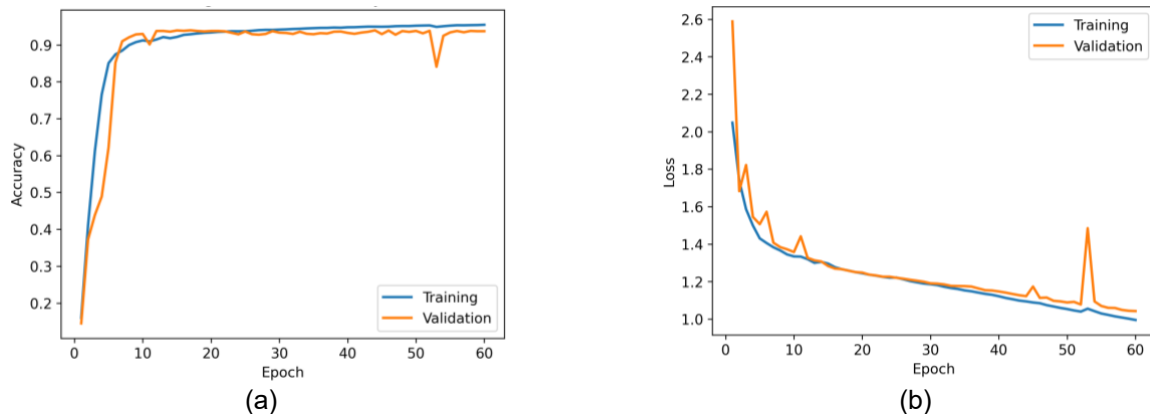


Fig. 12. Training performance of the U-Net retinal vessel segmentation model: (a) training and validation accuracy and (b) training and validation loss.

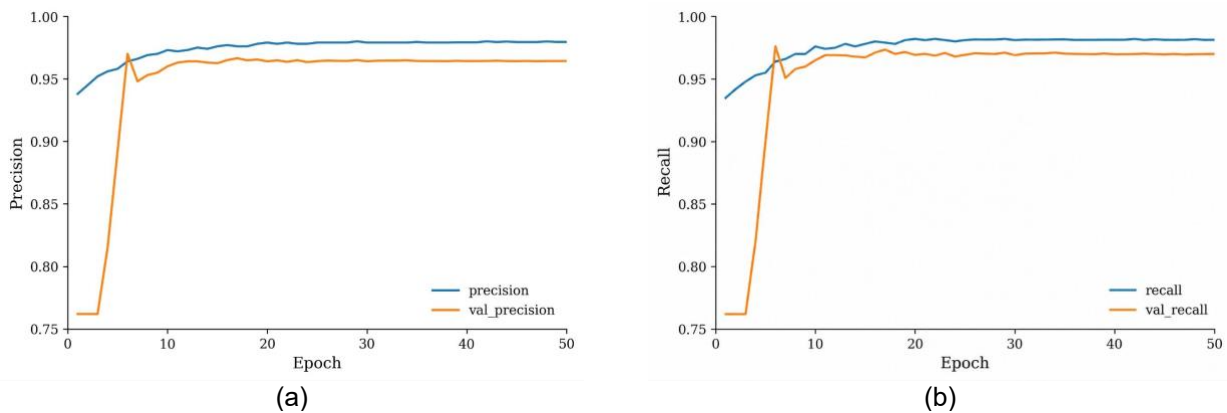


Fig. 11. Training performance of the U-Net retinal vessel segmentation model: (a) training and validation precision and (b) training and validation recall.

highest segmentation performance, and the training curves and visual examples show that U-Net can extract the main retinal blood vessel structures consistently.

D. Classification Result

The DR classification process was performed using the DLSTM-121 architecture with segmented retinal images obtained from U-Net. Before segmentation, PCA was

layer. The combination of DenseNet-121 and LSTM forms the DLSTM-121 architecture. The model was trained for 100 epochs with 938 iterations per epoch, and the optimal weights were used for testing. Fig. 14 presents the training process results. Based on Fig. 14a, the training accuracy increased rapidly during the early epochs and stabilized at approximately 0.98, while the

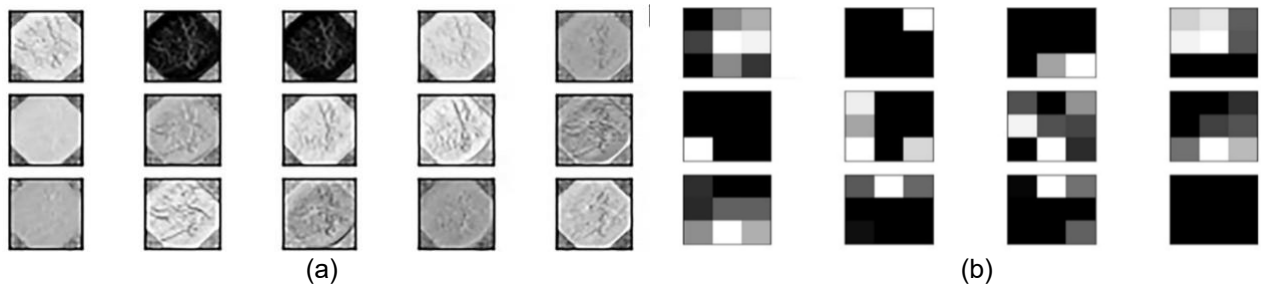


Fig. 13. Examples of intermediate feature representations in the DR classification process (a) convolution first-layer output and (b) concatenation layer output in the first dense block.

validation accuracy stabilized at approximately 0.97. The gap between the two curves was about 0.01, indicating consistent learning between the training and validation data. As shown in Fig. 14b, the training loss decreased to below 0.05 within the first 10 epochs and remained close to zero, while the validation loss also decreased and remained low with minor fluctuations. The relatively small difference between the training and validation curves suggests that no substantial overfitting was observed during training.

After training, the best weights from each run were used to evaluate the proposed DLSTM-121 model on 13 original retinal images, consisting of 10 DR and 3 Normal images, across five independent runs. The model

IV. Discussion

An ablation study was conducted to evaluate the effect of the classification model and PCA-based dimensionality reduction before U-Net segmentation. Five configurations were evaluated under the same experimental settings: LSTM with U-Net, DenseNet-121 with U-Net, DLSTM-121 with U-Net, DenseNet-121 with U-Net using PCA75 images, and DLSTM-121 with U-Net using PCA75 images.

Table 5 shows that the LSTM with U-Net configuration achieved the lowest overall classification performance, whereas DenseNet-121 and DLSTM-121

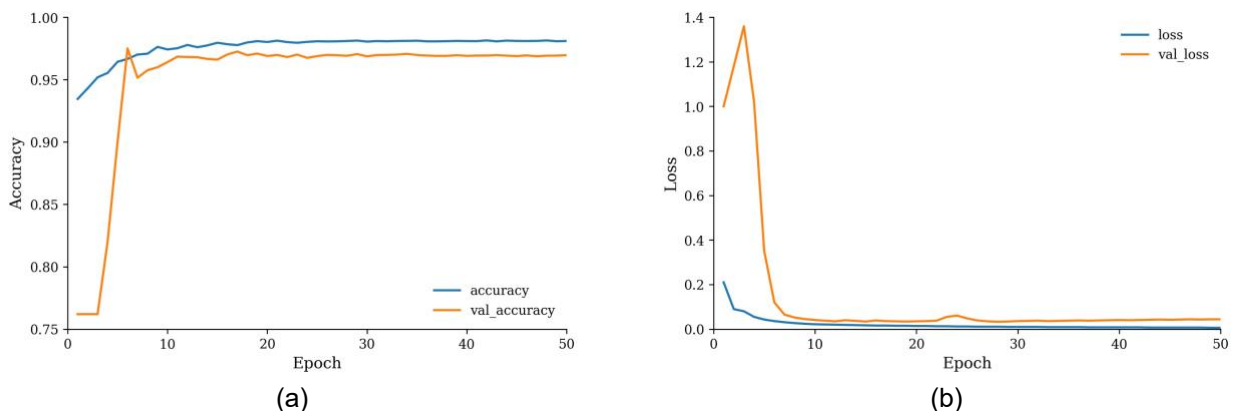


Fig. 14. Training performance of the DLSTM-121 diabetic retinopathy classification model: (a) training and validation accuracy and (b) training and validation loss.

correctly classified all DR images in every run, resulting in TP of 10.00 ± 0.00 and FN of 0.00 ± 0.00 . For Normal images, the mean TN and FP were 2.60 ± 0.89 and 0.40 ± 0.89 , respectively. These decimal values represent the average results across the five runs. Overall, the model achieved an accuracy of $96.92\% \pm 6.88\%$, sensitivity of $100.00\% \pm 0.00\%$, specificity of $86.67\% \pm 29.81\%$, precision of $96.67\% \pm 7.45\%$, F1-score of $98.18\% \pm 4.07\%$, and AUC of $99.33\% \pm 1.49\%$. The ROC curve moves up until it approaches 1; thus, the area under the ROC curve (AUC) increases and reaches 0.99, as shown in Fig. 15.

with U-Net achieved higher performance. The Applying PCA75 before U-Net segmentation further improved classification performance. DLSTM-121 with PCA75 achieved the highest accuracy (96.92%) and 100% sensitivity. Although DenseNet-121 with U-Net without PCA achieved the highest specificity (100%), its sensitivity was lower. DLSTM-121 with PCA75 also achieved a higher specificity than DenseNet-121 with PCA75 and a higher F1-score, indicating a better balance between DR detection and Normal-image classification. For AUC, DenseNet-121 with PCA75 achieved 100%, slightly higher than the 99.33%

obtained by DLSTM-121. This difference indicates that DenseNet-121 provided slightly better class separability across thresholds, whereas DLSTM-121

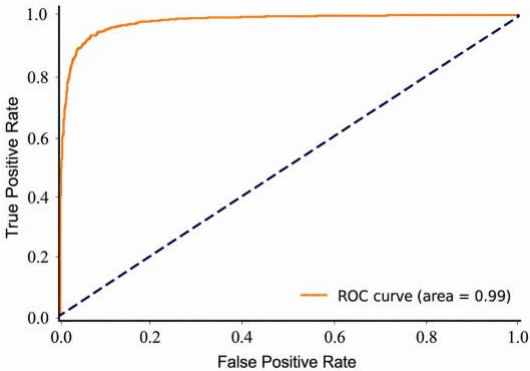


Fig. 15. Receiver operating characteristic (ROC) curve of the proposed PCA75-DLSTM-121 model for diabetic retinopathy classification.

provided better performance at the selected classification threshold based on accuracy, sensitivity,

HOG-ELM [48] achieved the lowest sensitivity of 75.95% on IDRiD. The substantial variation in sensitivity across studies indicates that DR detection performance depends not only on the classification model but also on dataset characteristics and experimental conditions. The 100% sensitivity obtained by PCA-DLSTM121 therefore demonstrates strong DR detection performance on the evaluated STARE test set, which is particularly relevant to automated DR screening. PCA-DLSTM121 also achieved the highest reported F1-score among the methods with available F1-score results, at 98.18%, reflecting a favorable balance between precision and sensitivity. In contrast, PCA-ResNet152 [49] reported the lowest F1-score of 72.87%. The variation in F1-score further indicates that differences in datasets, class definitions, and evaluation procedures contribute to the differences in reported classification performance. In terms of accuracy, PCA-DLSTM121 achieved 96.92%, while the highest accuracy was reported by PCA-BoV SVM [50] at 99.40%. Although the proposed method did not achieve the highest accuracy, its 96.92% accuracy was accompanied by a sensitivity of 100%, indicating that

Table 5. Ablation Study of DenseNet-121, LSTM, and PCA Components for Diabetic Retinopathy Classification.

Model Classification	Reduction and Segmentation	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)	AUC (%)
LSTM	U-Net	61.54	66	46.67	70.47	62
DenseNet-121	U-Net	83.08	78	100	87.58	95.33
DLSTM-121	U-Net	81.54	80	86.67	87.02	86
DenseNet-121	PCA75-UNet	92.31	100	66.67	95.24	100
DLSTM-121	PCA75-UNet	96.92	100	86.67	98.18	99.33

specificity, and F1-score. Overall, using PCA75 before U-Net segmentation followed by DLSTM-121 classification provided the best overall performance among the evaluated configurations. To compare the proposed approach with previous studies, Table 6 presents accuracy (Acc), sensitivity (Sn), specificity (Sp), and F1-score. The compared studies used different datasets, disease class definitions, sample sizes, image distributions, preprocessing procedures, sample independence, classification tasks, and test protocols. The proposed method was evaluated for binary classification of DR and Non-DR using the STARE dataset, with 15,000 augmented samples generated from 103 original training images. PCA-DLSTM121 achieved an accuracy of 96.92%, sensitivity of 100%, specificity of 86.67%, and F1-score of 98.18%. The sensitivity of PCA-DLSTM121 was the highest reported among the methods presented in Table 6, indicating that all DR cases in the evaluated test set were correctly identified. In comparison, PCA-

the model achieved particularly strong detection of DR cases. The specificity of 86.67% indicates that the model was also able to identify Non-DR cases, although its performance in this class was comparatively lower than its DR detection performance. This performance profile shows that PCA-DLSTM121 achieved strong DR detection while maintaining high accuracy and F1-score.

The comparison further demonstrates that the same classification method can produce substantially different results across datasets. For example, PCA-HOG-ELM [48] achieved 92.21% accuracy and 93.42% sensitivity on APTOS, but these values decreased to 72.12% and 75.95%, respectively, on IDRiD. The substantial decline in both accuracy and sensitivity for the same method across these datasets demonstrates that the reported performance is strongly dependent on the dataset on which the model is evaluated, rather than being determined by the model architecture alone. Differences in image distributions, disease

Table 6. Performance Comparison of the Proposed PCA75-DLSTM-121 Model with Previous Diabetic Retinopathy Classification Methods.

Method	Years	Number of Samples	Dataset	Labels	Acc (%)	Sn (%)	Sp (%)	F1-Score (%)
PCA-VGG19 [22]	2018	35,126	Kaggle Diabetic Retinopathy Dataset	5 labels (Non-Diabetic Retinopathy, Mild, Moderate, Severe, & Proliferative Diabetic Retinopathy)	92.21	-	-	-
PCA-MLP [49]	2019	128	Structured Analysis of the Retina	2 labels (Diabetic R & Non-Diabetic Retinopathy)	86.38	-	-	-
PCA-DNN [23]	2020	-	Methods to Evaluate Segmentation and Indexing Techniques in the Field of Retinal Ophthalmology	2 labels (Diabetic Retinopathy & Non-Diabetic Retinopathy)	90.07	90.4	94.3	-
PCA-HOG ELM [48]	2022	950	Asia Pacific Tele-Ophthalmology Society Blindness Detection Dataset	2 labels (Diabetic Retinopathy & Non-Diabetic Retinopathy)	92.21	93.42	77.27	95.3
PCA-HOG ELM [48]	2022	950	Indian Diabetic Retinopathy Image Dataset	2 labels (Diabetic Retinopathy & Non-Diabetic Retinopathy)	72.12	75.95	60	80.54
PCA-Resnet152 [51]	2023	1445	Diabetic Retinopathy and Clinically Significant Macular Edema Dataset	6 labels (No Diabetic Retinopathy, Mild Non-Proliferative Diabetic Retinopathy, Moderate Non-Proliferative Diabetic Retinopathy, Severe Non-Proliferative Diabetic Retinopathy, Early Proliferative Diabetic Retinopathy, and high-risk Proliferative Diabetic Retinopathy)	94.4	76.35	-	72.87
PCA-EfficientNetB3 [52]	2026	700	Methods to Evaluate Segmentation and Indexing Techniques in the Field of Retinal Ophthalmology	2 labels (Diabetic Retinopathy & Non Diabetic Retinopathy)	97.03	-	93.33	-
PCA-BoV SVM [50]	2025	130	Diabetic Retinopathy Database 0	3 labels (Mild Non-Proliferative Diabetic Retinopathy, Moderate Non-Proliferative Diabetic Retinopathy, Severe Non-Proliferative Diabetic Retinopathy)	99.4	99.75	98.7	-
Proposed Method	2026	15000	Structured Analysis of the Retina	2 labels (Diabetic Retinopathy & Non Diabetic Retinopathy)	96.92	100	86.6	98.1

Description: The bold value is the highest percentage

protocols may contribute to the observed variation. Therefore, the performance differences reported in Table 6 reflect not only differences in model architecture but also differences in the data and experimental conditions under which each model was evaluated. Regarding sample size, PCA-DLSTM121 used 15,000 augmented samples, which was larger than several studies but smaller than the 35,126 samples used by PCA-VGG19 [22]. However, these 15,000 samples were generated from 103 original training images and therefore do not represent an equivalent number of independent retinal images. This distinction is important because the number of augmented samples reflects the amount of data available for model training, whereas the number of original images reflects the diversity and independence of the underlying retinal data.

Another important aspect of the proposed method is that PCA was fitted only on the training data, and the resulting parameters were subsequently applied to the validation and testing data. This procedure prevents evaluation data from contributing to the determination of the principal components, thereby reducing the risk of data leakage. Although previous studies also applied PCA, their methodological descriptions do not always clearly specify whether PCA was fitted exclusively on the training data. Thus, differences in PCA processing procedures represent another methodological factor that may contribute to variation in the reported results. Evidence of the model's generalization capability is currently limited to the STARE dataset. Although the training and validation curves did not show an obvious indication of overfitting, these results do not establish how the model would perform on independent retinal datasets. The observed performance therefore primarily reflects the characteristics of STARE, while robustness across different datasets remains unassessed. External validation using independent datasets would provide stronger evidence of the robustness of the retinal model's generalization capability across different datasets.

Several limitations should also be noted. The limited number of original and independent test images may affect the stability of performance estimates, while the use of a single dataset limits assessment across different retinal image distributions. In addition, although augmentation increased the training data to 15,000 samples, the underlying data remained derived from a limited number of original images and did not represent 15,000 independent retinal images. Future studies should evaluate the proposed method using larger and more diverse datasets, independent external test sets, and subject-level validation to further assess its robustness and generalization.

V. Conclusion

This study aimed to develop a PCA-DLSTM-121-based method for diabetic retinopathy (DR) classification using retinal fundus images from the STARE dataset. The proposed method combined data augmentation, PCA, and the DLSTM-121 architecture. Data augmentation was applied to increase training data diversity, while PCA was fitted only on the training data, and the resulting transformation was applied to the validation and test data to reduce potential data leakage. The proposed PCA-DLSTM-121 achieved an accuracy of 96.92%, sensitivity of 100%, specificity of 86.67%, and F1-score of 98.18%, indicating promising performance in identifying DR cases. However, the specificity of 86.67% indicates that the identification of Non-DR cases remains an area for improvement. The training and validation curves did not show an obvious indication of overfitting on the evaluated dataset. These results should be interpreted with caution because the test set consisted of only 13 original images. In addition, external validation was not performed, so the current findings do not establish the generalization capability of the proposed method to other retinal datasets or populations. Future studies should perform external validation using independent retinal datasets or subject-level cross-validation to further assess the robustness and generalizability of the proposed method. Further evaluation using retinal images with different image qualities, illumination conditions, and population characteristics is also needed. Improving the identification of Non-DR cases while maintaining high sensitivity remains an important direction for future research.

Acknowledgment

The authors acknowledge Politeknik Negeri Lhokseumawe for institutional support and research facilities provided during this study.

Funding

This research was supported by Politeknik Negeri Lhokseumawe under Innovation Research Grant No. 062/SK.LPPM-01/2025.

Data Availability

The dataset used in this study is publicly available from the STARE database. <https://cecas.clemson.edu/~aho/over/stare/>.

Author Contribution

Muhammad Arhami: Conceptualization, Methodology, Formal Analysis, Writing – Original Draft, Supervision. Marwan Ramli: Software, Data Curation, Validation,

Writing – Original Draft. Melinda: Investigation, Data Curation, Visualization. Anita Desiani: Methodology, Validation, Writing – Review & Editing. Husaini: Resources, Project Administration, Investigation. Musta'inul Abdi: Formal Analysis, Validation, Writing – Review & Editing.

Declarations

Declaration of Generative Artificial Intelligence Use.

Generative artificial intelligence tools were used solely to assist with language editing and grammatical improvement. The authors reviewed and verified all scientific content, interpretations, analyses, and conclusions and take full responsibility for the final manuscript.

Ethical Approval

Not applicable. This study used a publicly available retinal image dataset and did not involve direct recruitment of human participants.

Consent for Participants.

Not applicable.

Consent for Publication.

Not applicable because the study used publicly available and de-identified image data.

Competing Interests

The authors declare that they have no competing interests.

References

- [1] J. Zhou and B. Chen, "Retinal Cell Damage in Diabetic Retinopathy," *cells*, vol. 12, no. 9, p. 1342, 2023, doi: 10.3390/cells12091342.
- [2] (WHO) World Health Organization, "Global Report on Diabetes," Geneva, Switzerland, 2016. [Online]. Available: <https://www.who.int/publications/i/item/9789241565257>
- [3] T. R. Rao and G. N. Tejomurtulada, "Diabetes Mellitus: A Review," *Int. J. Med. Sci. Pharma Res.*, vol. 10, no. 2, pp. 5–9, 2024, doi: 10.22270/ijmspr.v10i2.97.
- [4] F. K. Malerbi *et al.*, "Automated Identification of Different Severity Levels of Diabetic Retinopathy Using a Handheld Fundus Camera and Single-Image Protocol," *Ophthalmol. Sci.*, vol. 4, no. 4, p. 100481, 2024, doi: 10.1016/j.xops.2024.100481.
- [5] Y. Kim, A. F. M. S. Uddin, and S. H. Bae, "Local Augment: Utilizing Local Bias Property of Convolutional Neural Networks for Data Augmentation," *IEEE Access*, vol. 9, pp. 15191–15199, 2021, doi: 10.1109/ACCESS.2021.3050758.
- [6] A. Desiani, Erwin, B. Suprihatin, S. Yahdin, A. I. Putri, and F. R. Husein, "Bi-path Architecture of CNN Segmentation and Classification Method for Cervical Cancer Disorders Based on Pap-smear Images," *IAENG Int. J. Comput. Sci.*, vol. 48, no. 3, pp. 782–791, 2021, [Online]. Available: https://www.iaeng.org/IJCS/issues_v48/issue_3/IJCS_48_3_37.pdf
- [7] J. Gupta, S. Pathak, and G. Kumar, "Deep Learning (CNN) and Transfer Learning: A Review," *J. Phys. Conf. Ser.*, vol. 2273, no. 1, p. 012029, 2022, doi: 10.1088/1742-6596/2273/1/012029.
- [8] N. Hasan, Y. Bao, A. Shawon, and Y. Huang, "DenseNet Convolutional Neural Networks Application for Predicting COVID-19 Using CT Image," *SN Comput. Sci.*, vol. 2, no. 5, p. 389, 2021, doi: <https://doi.org/10.1007/s42979-021-00782-7>.
- [9] C. Mohanty *et al.*, "Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy," *Sensors*, vol. 23, no. 12, p. 5726, 2023, doi: 10.3390/s23125726.
- [10] S. Das and S. Majumder, "Detection of diabetic retinopathy 5-stages by denseNet, a CNN model," *Franklin Open*, vol. 15, pp. 1–8, 2026, doi: 10.1016/j.fraope.2026.100577.
- [11] G. Alwakid, W. Gouda, M. Humayun, and N. Z. Jhanjhi, "Deep learning-enhanced diabetic retinopathy image classification," *Digit. Heal.*, vol. 9, p. 20552076231194944, 2023, doi: 10.1177/20552076231194942.
- [12] K. Zhang, Y. Guo, X. Wang, J. Yuan, and Q. Ding, "Multiple Feature Reweight DenseNet for Image Classification," *IEEE Access*, vol. 7, pp. 9872–9880, 2019, doi: 10.1109/ACCESS.2018.2890127.
- [13] S. M. Al-selwi, M. F. Hassan, S. J. Abdulkadir, and A. Muneer, "LSTM Inefficiency in Long-Term Dependencies Regression Problems," *J. Adv. Res. Appl. Sci. Eng. Technol.*, vol. 30, no. 3, pp. 16–31, 2023, doi: 10.37934/araset.30.3.1631.
- [14] K. Glynis, Z. Kapelan, M. Bakker, and R. Taormina, "Leveraging Transfer Learning in LSTM Neural Networks for Data - Efficient Burst Detection in Water Distribution Systems," *Water Resour. Manag.*, vol. 37, no. 15, pp. 5953–5972, 2023, doi: 10.1007/s11269-023-03637-3.
- [15] T. D. Pham, "Time-frequency time-space LSTM for robust classification of physiological signals," *Sci. Rep.*, vol. 11, no. 1, p. 6936, 2021,

- doi: 10.1038/s41598-021-86432-7.
- [16] B. M. S. Rani, V. R. Ratna, V. P. Srinivasan, S. Thenmalar, and R. Kanimozhi, "Disease prediction based retinal segmentation using bi-directional ConvLSTMU-Net," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, pp. 1–10, 2021, doi: 10.1007/s12652-021-03561-5.
- [17] R. Vaydande, "Retinal Fundus Image Classification using LSTM - Convolution Neural Network," National College of Ireland, Ireland, 2022. [Online]. Available: <https://norma.ncirl.ie/6332/>
- [18] T. Palaniswamy and M. Vellingiri, "Internet of Things and Deep Learning Enabled Diabetic Retinopathy Diagnosis Using Retinal Fundus Images," *IEEE Access*, vol. 11, pp. 27590–27601, 2023, doi: 10.1109/ACCESS.2023.3257988.
- [19] H. A. H. Mahmoud, A. A. AlArfaj, and A. M. Hafez, "A Fast Hybrid Classification Algorithm with Feature Reduction for Medical Images," *Appl. Bionics Biomech.*, vol. 2022, pp. 1–11, Jan. 2022, doi: 10.1155/2022/1367366.
- [20] M. R. Vinutha, J. Chandrika, B. Krishnan, and S. A. Kokatnoor, "EPCA—Enhanced Principal Component Analysis for Medical Data Dimensionality Reduction," *SN Comput. Sci.*, vol. 4, no. 3, p. 243, 2023, doi: 10.1007/s42979-023-01677-5.
- [21] F. Saeed, M. Hussain, and H. A. Aboalsamh, "Automatic Diabetic Retinopathy Diagnosis Using Adaptive Fine-Tuned Convolutional Neural Network," *IEEE Access*, vol. 9, pp. 41344–41359, 2021, doi: 10.1109/ACCESS.2021.3065273.
- [22] M. Mateen, J. Wen, Nasrullah, S. Song, and Z. Huang, "Fundus Image Classification Using VGG-19 Architecture with PCA and SVD," *Symmetry (Basel)*, vol. 11, no. 1, p. 1, 2018, doi: 10.3390/sym11010001.
- [23] T. R. Gadekallu *et al.*, "Early Detection of Diabetic Retinopathy Using PCA-Firefly Based Deep Learning Model," *Electron.*, vol. 9, no. 2, pp. 1–16, 2020, doi: 10.3390/electronics9020274.
- [24] R. F. Mansour, "Deep-Learning-Based Automatic Computer-Aided Diagnosis System for Diabetic Retinopathy," *Biomed. Eng. Lett.*, vol. 8, no. 1, pp. 41–57, 2018, doi: 10.1007/s13534-017-0047-y.
- [25] M. Mateen, J. Wen, M. Hassan, N. Nasrullah, S. Sun, and S. Hayat, "Automatic Detection of Diabetic Retinopathy: A Review on Datasets, Methods and Evaluation Metrics," *IEEE Access*, vol. 8, pp. 48784–48811, 2020, doi: 10.1109/ACCESS.2020.2980055.
- [26] A. Imran, J. Li, Y. Pei, J. Yang, and Q. Wang, "Comparative Analysis of Vessel Segmentation Techniques in Retinal Images," *IEEE Access*, vol. 7, pp. 114862–114887, 2019, doi: 10.1109/ACCESS.2019.2935912.
- [27] H. G. T. Özsoy, E. E. Şen, F. E. Urfalı, G. A. Aydın, B. Akyüz, and E. Dandil, "Measurement-Oriented Evaluation of Deep Learning Models for Automated Fetal Head Biometry in Prenatal Ultrasound," *J. Clin. Med.*, vol. 15, no. 18, p. 7127, 2026, doi: 10.3390/jcm15187127.
- [28] M. Hayati *et al.*, "Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning," in *7th International Conference on Computer Science and Computational Intelligence 2022*, Elsevier B.V., 2023, pp. 57–66. doi: 10.1016/j.procs.2022.12.111.
- [29] J. P. Bharadiya, "A Tutorial on Principal Component Analysis for Dimensionality Reduction in Machine Learning," *Int. J. Innov. Sci. Res. Technol.*, vol. 8, no. 5, pp. 2028–2032, 2023, doi: 10.5281/zenodo.8002436.
- [30] W. Ding, "SVM-Based Feature Selection for Differential Space Fusion and Its Application to Diabetic Fundus Image Classification," *IEEE Access*, vol. 7, pp. 149493–149502, 2019, doi: 10.1109/ACCESS.2019.2944899.
- [31] H. Zhao, J. Zheng, J. Xu, and W. Deng, "Fault Diagnosis Method Based on Principal Component Analysis and Broad Learning System," *IEEE Access*, vol. 7[1], pp. 99263–99272, 2019, doi: 10.1109/ACCESS.2019.2929094.
- [32] Y. Zhang, B. Zhang, and Z. Wu, "Multi-Model Modeling of CFB Boiler Bed Temperature System Based on Principal Component Analysis," *IEEE Access*, vol. 8, pp. 389–399, 2020, doi: 10.1109/ACCESS.2019.2961414.
- [33] G. T. Reddy *et al.*, "Analysis of Dimensionality Reduction Techniques on Big Data," *IEEE Access*, vol. 8, pp. 54776–54788, 2020, doi: 10.1109/ACCESS.2020.2980942.
- [34] T. Tokuda, B. Goodrich, A. Gelman, and F. Tuerlinckx, "Visualizing Distributions of Covariance Matrices," *J. Data Sci. Stat. Vis.*, vol. 5, no. 7, pp. 1–30, 2025, doi: 10.52933/jdssv.v5i7.132.
- [35] T. A. Soomro *et al.*, "Deep learning models for retinal blood vessels segmentation: a review," *IEEE Access*, vol. 7, pp. 71696–71717, 2019, doi: 10.1109/ACCESS.2019.2920616.
- [36] M. Mubashar, H. Ali, C. Gronlund, and S. Azmat, "R2U++ : a multiscale recurrent residual

- U-Net with dense skip connections for medical image segmentation," *Neural Comput. Appl.*, vol. 34, no. 20, pp. 17723–17739, 2022, doi: 10.1007/s00521-022-07419-7.
- [37] M. H. Reza *et al.*, "A comprehensive review of convolutional neural networks: foundations, enhancements and applications," *Neural Comput. Appl.*, vol. 38, p. 56, 2026, doi: 10.1007/s00521-025-11827-w.
- [38] R. Balestrierio and R. G. Baraniuk, "Batch Normalization Explained," *ArXiv Prepr.*, 2022, [Online]. Available: <http://10.0.189.166/arXiv.2209.14778>
- [39] Y. Yuan and Y. Cheng, "Medical image segmentation with UNet-based multi-scale context fusion," *Sci. Rep.*, vol. 14, p. 15687, 2024, doi: 10.1038/s41598-024-66585-x.
- [40] I. Kandel and M. Castelli, "Transfer Learning with Convolutional Neural Networks for Diabetic Retinopathy Image Classification. A Review," *Appl. Sci.*, vol. 10, no. 6, p. 2021, 2020, doi: 10.3390/app10062021.
- [41] S. Albahli, T. Nazir, A. Irtaza, and A. Javed, "Recognition and detection of diabetic retinopathy using densenet-65 based faster-RCNN," *Comput. Mater. Contin.*, vol. 67, no. 2, pp. 1333–1351, 2021, doi: 10.32604/cmc.2021.014691.
- [42] W. Hou, "Statistical and machine learning methods for multi-step earthquake frequency forecasting in Indonesian regions," *Nat. Hazards*, vol. 122, no. 40, pp. 1–21, 2026, doi: 10.1007/s11069-025-07744-9.
- [43] C. Vasile, A. Ulm, and C. Bira, "Image Processing Hardware Acceleration — A Review of Operations Involved and Current Hardware Approaches," *J. Imaging*, vol. 10, no. 12, p. 298, 2024, doi: 10.3390/jimaging10120298.
- [44] Y. Wang, Y. Zhao, and S. Addepalli, "Practical Options for Adopting Recurrent Neural Network and Its Variants on Remaining Useful Life Prediction," *Chinese J. Mech. Eng.*, vol. 34, p. 69, 2021, doi: 10.1038/s41598-024-66585-x.
- [45] A. Desiani *et al.*, "Weighted Voting Ensemble for Enhanced Diabetic Retinopathy Classification using CNN Architectures," *J. Infotel*, vol. 16, no. 1, pp. 136–155, 2024, doi: 10.20895/infotel.v16i1.999.
- [46] D. Nagpal, S. N. Panda, M. Malarvel, P. A. Pattanaik, and M. Z. Khan, "A review of diabetic retinopathy: Datasets, approaches, evaluation metrics and future trends," *J. King Saud Univ. – Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7138–7152, 2022, doi: 10.1016/j.jksuci.2021.06.006.
- [47] V. Raudonis *et al.*, "Automatic Detection of Microaneurysms in Fundus Images Using an Ensemble-Based Segmentation Method," *Sensors*, vol. 23, no. 7, p. 3431, 2023, doi: 10.3390/s23073431.
- [48] M. A. A. Albadr, M. Ayob, S. Tiun, F. T. AL-Dhief, and M. K. Hasan, "Gray wolf optimization-extreme learning machine approach for diabetic retinopathy detection," *Front. Public Heal.*, vol. 10, p. 925901, 2022, doi: 10.3389/fpubh.2022.925901.
- [49] R. Arasy and Basari, "Detection of hypertensive retinopathy using principal component analysis (PCA) and backpropagation neural network methods," in *3rd International Symposium of Biomedical Engineering's Recent Progress in Biomaterials, Drugs Development, and Medical Devices, ISBE 2018*, P. K. Praswasti P.D.K. Wulan, Misri Gozan, Sotya Astutiningsih, Ghiska Ramahdita, Radon Dhelika, Ed., Jakarta, Indonesia: American Institute of Physics Inc., 2019, p. 040002. doi: 10.1063/1.5096735.
- [50] G. Vinuja, R. Devi, N. Salm, A. Niyaz, and R. Kamnani, "Improved Detection of Early Diabetic Retinopathy using Contrast Limited Adaptive Histogram Equalization and Principal Component Analysis with Bag of Visual Words based SVM," in *2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE)*, IEEE, 2025, pp. 1–8. doi: 10.1109/ICICKE65317.2025.11136603.
- [51] T. M. Usman, Y. K. Saheed, D. Ignace, and A. Nsang, "Diabetic retinopathy detection using principal component analysis multi-label feature extraction and classification," *Int. J. Cogn. Comput. Eng.*, vol. 4, pp. 78–88, 2023, doi: 10.1016/j.ijcce.2023.02.002.
- [52] L. Costaner, Lisnawita, and Guntoro, "Diabetic Retinopathy Detection via PCA-MLP Using Independent GLCM and EfficientNetB3 Descriptors," *J. Teknol. Inf. dan Komun.*, vol. 17, no. 1, pp. 37–48, 2026, doi: 10.31849/digitalzone.v17i1.33099.

Author Biography



Muhammad Arhami completed his bachelor's degree at Universitas Syiah Kuala in 1998, followed by a master's degree at Universitas Gadjah Mada in 2004. Currently, he is affiliated with Politeknik Negeri Lhokseumawe, specializing in Informatics Engineering. He is also a doctoral student at Universitas Syiahkuala in 2025. His current research interests include machine learning,

artificial intelligence, data mining, natural language processing, and expert systems. Arhami's dedication to advancing these fields contributes significantly to research and education in technology and computer science.



Marwan Ramli is a Professor in Applied Mathematics and currently an active lecturer and researcher at the Department of Mathematics and the Mathematics Graduate Study Program, Syiah Kuala University. Prof. Marwan was born on November 25, 1971, in Tanjung Pinang, Kepulauan Riau. He completed his Ph.D. in Applied Mathematics, especially in Extreme Water Waves Generation, from the Bandung Institute of Technology, Indonesia, in collaboration with the University of Twente, The Netherlands, in 2006, under the supervision of Prof. Dr. M. Ansjar, Dr. Andonowati, Prof. Dr. Edy Soewono, and Prof. Dr. E. van Groesen. He finished his Master degree in Mathematics from the Bandung Institute of Technology, Indonesia, in 1999, under supervision Prof. Dr. Sunardi Wirjosudirjo. He finished his BSc in Mathematics from Syiah Kuala University, Indonesia, in 1996.



Melinda was born in Bireuen, Aceh, on June 10, 1979. She became a professor in the Department of Electrical and Computer Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh, in 2002. She completed her master's degree at the Faculty of Electrical Department, University of Southampton, United Kingdom, with a concentration in field study of Radio Frequency Communication Systems in 2009. She completed her Doctoral degree at the Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia, in February 2018. She has been with the Department of Electrical Engineering, Faculty of Engineering, Universitas Syiah Kuala since 2002. She is also a member of IEEE. Her research interests include multimedia signal processing and fluctuation processing. She can be contacted at email: melinda@usk.ac.id.



Anita Desiani; in 2020, she was working on a project for her Doctoral Program at the Faculty of Mathematics and Natural Sciences, Universitas Sriwijaya. She received a bachelor's degree in mathematics from Universitas Sriwijaya in 2000, a master's degree in Computer Science from Universitas Gadjah

Mada in 2003, and a doctoral degree in Mathematics and Science Department in 2022. In 2004, she joined the Mathematics Department at Universitas Sriwijaya as a lecturer, where she has remained until now. Her current research interests include the fields of data mining, image processing, pattern recognition, computer vision, and artificial intelligence.



Musta'inul Abdi was born in Rokan Hulu, Riau, on October 30, 1991. He received the Bachelor of Applied Science degree in Informatics Engineering from Politeknik Negeri Lhokseumawe, Indonesia, in 2014. He received a master's degree in a related field from Institut Teknologi Sepuluh Nopember in 2017. He is an assistant professor in the Department of Information and Computer Technology at Politeknik Negeri Lhokseumawe. His research interests include artificial intelligence, large language models, deep learning, and computer vision.



Husaini was born in Sigli, Aceh, on October 31, 1973. He received the B.S. degree in Mathematics from Universitas Syiah Kuala, Banda Aceh, Indonesia, in 1998, and an M.IT. degree in Information Technology from Universiti Kebangsaan Malaysia in 2009. He is a Senior Lecturer in the Department of Information Technology and Computer at Politeknik Negeri Lhokseumawe, Indonesia. His research interests include the Internet of Things, artificial intelligence, and network security.