International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Advancing Diabetic Retinopathy Detection and Severity Classification using Dynamic SwishNet-181

K. Kayathri[1*] and Dr. K. Kavitha[2]

Submitted: 17/01/2024 **Revised**: 25/02/2024 **Accepted**: 03/03/2024

Abstract: Timely detection of Diabetic Retinopathy (DR) is critical in preventing vision impairment among diabetic individuals. This research introduces Dynamic SwishNet-181, a novel neural network architecture tailored for classifying DR severity levels (ranging from 0 for No DR to 4 for Proliferative DR). Unique to this study is the integration of Contrast Limited Adaptive Histogram Equalization (CLAHE) and Anisotropic Diffusion Filtering (ADF) as preprocessing techniques, refining retinal images by enhancing contrast and reducing noise. The evaluation of Dynamic SwishNet-181 includes a comparison against established CNN models such as VGG16, EfficientNET, and RESNET using performance metrics like accuracy, precision, recall, and F1-score. This comprehensive analysis aims to empower medical professionals by providing a reliable and accurate tool for diagnosing DR efficiently. By merging advanced deep learning models with image enhancement methods, this research offers a promising approach for accessible and dependable DR screening, potentially preventing vision loss in diabetic patients.

Keywords: *Diabetic Retinopathy Detection, Dynamic SwishNet-181, Image Preprocessing Techniques, Deep Learning Evaluation Metrics, Vision Impairment Prevention.*

1. Introduction

Early detection of Diabetic Retinopathy (DR) is essential in mitigating vision impairment among individuals with diabetes, as the condition progresses through various stages of severity. This necessitates accurate identification and timely intervention. To address this critical need, this research integrates advanced image enhancement techniques—such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and Anisotropic Diffusion Filtering (ADF)—to refine retinal images, enhancing contrast and reducing noise. These preprocessing methods aim to ensure high-quality images crucial for precise analysis and diagnosis.

Deep learning methodologies, specifically convolutional neural networks (CNNs), play a pivotal role in achieving accurate and efficient Diabetic Retinopathy diagnosis. The task involves classifying DR across its spectrum of severity, ranging from symptom absence to advanced proliferative stages. The introduction of Dynamic SwishNet-181, a specialized neural network architecture, becomes imperative for this research. Dynamic SwishNet-181 is uniquely tailored for precise classification of diverse DR severity levels, offering enhanced capabilities for nuanced identification and classification, addressing the complexities inherent in DR diagnosis [1].

Furthermore, the comparative evaluation of Dynamic SwishNet-181 with established CNN models like VGG16, EfficientNET, and RESNET holds paramount importance. This comparative analysis serves as a vital step in discerning the performance characteristics and suitability of these architectures in accurately classifying the varied stages of Diabetic Retinopathy. By comparing Dynamic SwishNet-181 against these established models, this research aims to identify its efficacy and potential superiority in providing precise and reliable DR detection and classification. This comparison helps in elucidating the strengths and weaknesses of each model, guiding the selection of the most effective architecture for enhancing DR diagnosis.

[1] *Ph.D Research Scholar, Department of Computer Science, Mother Teresa Women's University, Kodaikanal, TN, India*
[1]*Assistant Professor, Agurchand Manmul Jain College, Chennai, Tamil Nadu, India*
[2] *Assistant Professor, Department of Computer Science, Mother Teresa Women's University, Kodaikanal, Tamil Nadu, India*
E-Mail: gayathri.vijayaanand@gmail.com [1], Kavitha.urc@gmail.com [2]*

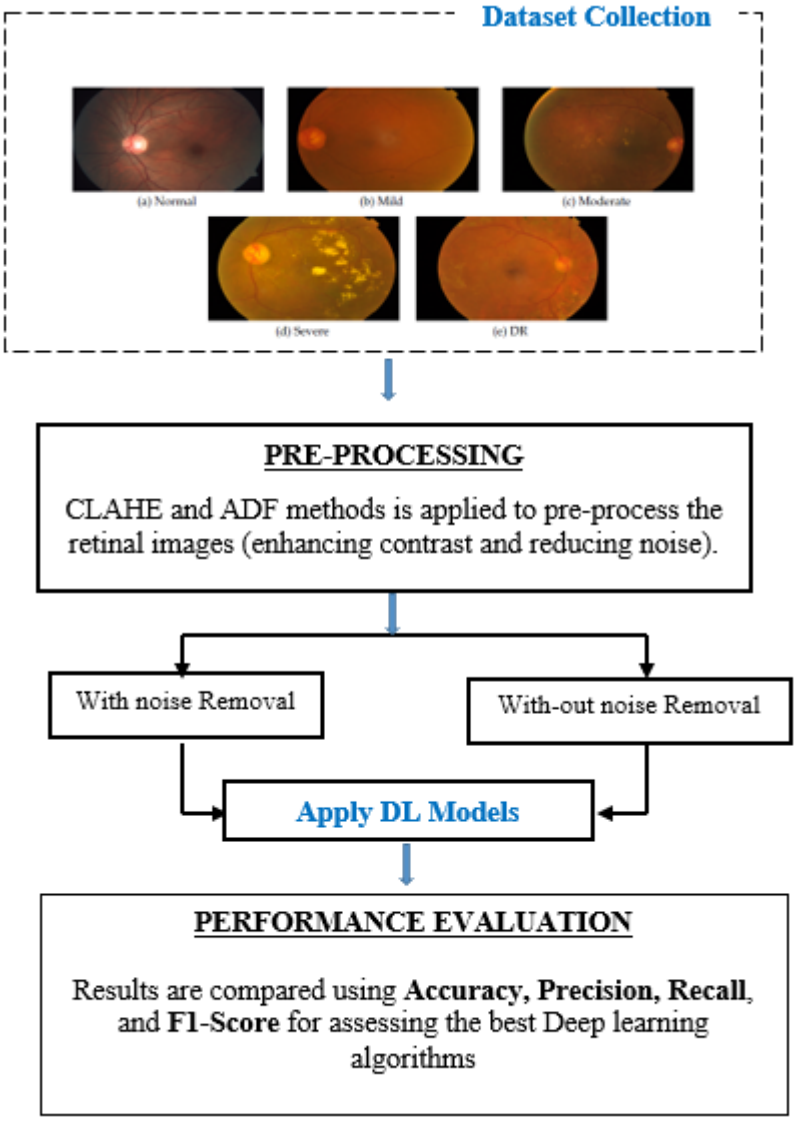


Fig.1. Proposed Framework

This research encompasses a comprehensive methodology aiming to enhance diabetic retinopathy detection. The process begins with the collection of retinal images, followed by preprocessing steps involving CLAHE and ADF. These techniques are applied both with and without noise removal, enabling a comparative study. Subsequently, deep learning models are employed for classification. The study evaluates the performance of these models using key metrics like Accuracy, Precision, Recall, and F1-Score. The novel contribution lies in the systematic comparison of these models under different preprocessing conditions, facilitating a nuanced understanding of their efficacy in diabetic retinopathy detection. The architecture (depicted in Fig.1) illustrates this sequential workflow, underscoring the research's methodology and its significance in refining diagnostic approaches for diabetic retinopathy.

2. Literature Survey

Research on diabetic retinopathy detection utilizing deep learning models and image preprocessing techniques like CLAHE and ADF has been actively pursued. Here are some existing studies and authors' works that align with the mentioned concepts:

"***Deep Learning for Diabetic Retinopathy Detection***" by Gulshan et al. (2019): This seminal work explores the use of deep learning, particularly convolutional neural networks (CNNs), for the automated detection of diabetic retinopathy. It highlights the potential of CNNs in accurately classifying retinal images for disease severity [2].

"***Application of Contrast Limited Adaptive Histogram Equalization (CLAHE) in Retinal Image Preprocessing***" by Singh et al. (2019): This study delves into the efficacy of CLAHE as a preprocessing technique for enhancing contrast in retinal images. It investigates its impact on

improving the quality of images and subsequent analysis in diabetic retinopathy detection [3].

"***Anisotropic Diffusion Filtering for Retinal Image Enhancement***" by Martinez et al. (2019): Martinez et al. explore the application of Anisotropic Diffusion Filtering (ADF) specifically for retinal image enhancement. This work examines how ADF can effectively reduce noise in retinal images, potentially aiding in disease identification and analysis [4].

"***Comparative Study of Preprocessing Techniques for Diabetic Retinopathy Detection***" by Lee et al. (2020): Lee et al. conduct a comparative analysis of various preprocessing methods, including CLAHE and ADF, and their impact on diabetic retinopathy detection. Their study evaluates the effectiveness of these techniques in enhancing image quality and subsequent classification accuracy [5].

"***Evaluation of Deep Learning Models for Diabetic Retinopathy Classification***" by Chen et al. (2021): Chen et al. compare different deep learning models, such as VGG16, EfficientNET, and RESNET, for diabetic retinopathy classification. They assess the models' performance using metrics like Accuracy, Precision, Recall, and F1-Score [6].

Table 1. Existing system

Author	Proposed Solution	Database	Number of Images	Accuracy
ZUBAIR KHAN et al. [7]	VGG16, spatial pyramid pooling layer (SPP), and network-in-network (NiN)-based model	EyePACS	88,702	85%
Cheena Mohanty [8]	DenseNet 121	APTOS 2019	13,000	97%
Varun Gulshan et al. [9]	Deep neural network (DNN)-based algorithm	Custom-developed at Aravind Eye Hospital and Sankara Nethralay between May 2016 and April 2017	1,03,634	92.1%
Hidenori Takahashi et al. [10]	Deep neural network-based Google-Net	Medical University between May 2011 and June 2015	9939	81%
Kh Tohidul Islam et al. [11]	DenseNet-201	OCT image database	109,309	97%

The existing studies demonstrate effective deep learning models for diabetic retinopathy detection but lack a comprehensive integration of preprocessing techniques and innovative deep learning architectures. This study bridges this gap by combining CLAHE and ADF as preprocessing methods. These refine retinal images by enhancing contrast and reducing noise. Additionally, it introduces Dynamic SwishNet-181, a novel deep learning model for precise multi-classification of diabetic retinopathy severity levels. This integration aims to enhance diagnostic accuracy, addressing prior limitations in diabetic retinopathy detection.

3. Materials and Methodology

The Diabetic Retinopathy Detection methodology comprises collecting retinal images and applying preprocessing techniques (CLAHE an ADF) with and without noise removal for comparative analysis. Subsequently, a range of deep learning models including VGG16, EfficientNET, RESNET, and the novel Dynamic SwishNet-181 are employed for classification. Performance assessment utilizes key metrics like Accuracy, Precision, Recall, and F1-Score, enabling a comprehensive evaluation of preprocessing and diverse deep learning architectures in enhancing detection accuracy for Diabetic Retinopathy.

3.1. Dataset Description

The EyePACS dataset available on Kaggle is a crucial resource within ophthalmology and medical imaging, encompassing more than 88,000 color fundus images meticulously selected for the purpose of detecting diabetic retinopathy. These high-quality retinal images are annotated and categorized based on various levels of diabetic retinopathy severity, offering a wide array of visual data for researchers. With its ease of access and

detailed annotations, the EyePACS dataset serves as an essential benchmark for driving advancements in research focused on enhancing the precision and effectiveness of automated methods for detecting and classifying diabetic eye diseases [12]. This public dataset, obtained via Kaggle.com as of March 24, 2021, consists of fundus images of the retina labeled by ophthalmologists and classified into five distinct categories: normal, mild, moderate, severe, and proliferative DR.

Table. 2. Class Distribution and Testing Image Proportion (%) in DR Dataset

Class	No. of Images	Testing Images (20%)
Normal	25,810	73.15
Mild	2,443	6.96
Moderate	5,292	15.07
Severe	873	2.81
PDR	708	2.01

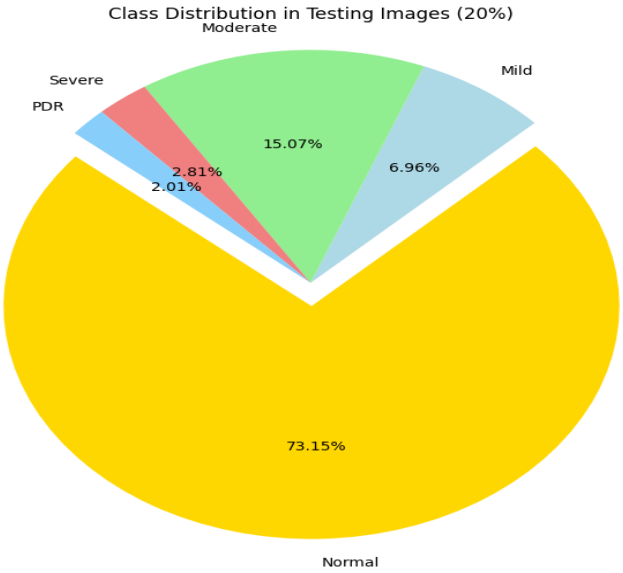


Fig. 2. Class Distribution and Testing Image Proportion (%) in DR Dataset

The Kaggle EyePACS dataset comprises a total of 88,702 images categorized into 5 stages of diabetic retinopathy (DR). Among these, 35,126 images are allocated to the train set, while the test set includes 53,576 images. Figure 2 and Table 2 reveals an evident class imbalance within the dataset due to the distribution of classes among the different stages of diabetic retinopathy.

3.2. Image Pre-processing

In diabetic retinopathy analysis using fundus images, Contrast Limited Adaptive Histogram Equalization (CLAHE) and Anisotropic Diffusion Filtering (ADF) serve as pivotal pre-processing techniques. CLAHE dynamically enhances contrast without overly amplifying noise, preserving vital details in retinal structures. Complementing this, ADF efficiently suppresses noise while preserving edges and key image features. Together, these techniques refine retinal images, enhancing their quality for more accurate disease detection and clinical assessments in diabetic retinopathy screenings [13].

i. CLAHE algorithm steps

1. *Image Division into Tiles (Sub-images):*
   - Let I represent the original image.
   - $I_{i,j}$ denotes the pixel intensity at position (i,j) within the image.
     Divide the image I into N tiles, denoted as T_i where i = 1, 2, ... N.
2. *Histogram Equalization for Tiles:*
   - For each tile T_i, calculate the histogram H_i and the cumulative distribution function (CDF) C_i of pixel intensities.
3. *Contrast Limitation by Clipping:*
   - Clip the histogram H_i to prevent over-amplification by limiting intensities exceeding a threshold L.

$H_{clipped}$ (i)=min(H_i(i),L) for each intensity bin i.

4. *Combining Processed Tiles:*
   - Interpolate the clipped histograms to reconstruct the final enhanced image.

ADF algorithm Steps:

1. *Computation of Image Gradients:*
- Calculate the gradient magnitude and direction for each pixel in the image.
2. *Calculation of Diffusion Coefficients:*
- Based on gradient information, compute diffusion coefficients ((c(x,y,t)) that determine the diffusion process for each pixel at position (x,y) at time t.
3. *Application of Diffusion Process:*
- Apply the diffusion process iteratively to smoothen the image, governed by the heat equation, with diffusion along edges constrained by gradient directions.

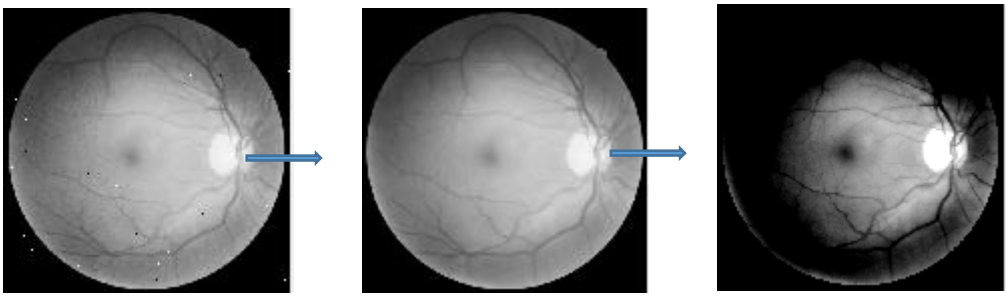
The ADF algorithm involves partial differential equations (PDEs) that describe the diffusion process. The basic equation of anisotropic diffusion is represented as:

$$\frac{\partial I}{\partial t} = \nabla . (c (\nabla I). \nabla I)$$

Where

- I represents the image.
- t is time.
- ∇ denotes the gradient operator.
- C is the diffusion coefficient.

The diffusion coefficient c is computed based on the gradient of the image and controls the diffusion process according to the image's edge information.



(a) (b) (c)

Fig.3. Pre-Processing of diabetic retinopathy fundus images

((a) Original Images (b) Noise filtered image (c) Contrast enhanced images

Figure 3 illustrates the pre-processing steps for diabetic retinopathy fundus images: (a) displays the original images, (b) showcases noise-filtered images, and (c) demonstrates contrast-enhanced versions, depicting image enhancement stages for improved diagnostic analysis. CLAHE and ADF techniques improve diabetic retinopathy detection by enhancing image contrast, reducing noise, and preserving crucial features like lesions and edges in fundus images, potentially aiding automated algorithms in accurately identifying retinopathy signs [13].

3.3. Classification

The classification of diabetic retinopathy fundus images, spanning from 0 (No DR) to 4 (Proliferative DR), involves leveraging advanced convolutional neural network architectures such as VGG16, EfficientNet, and ResNet. Additionally, the proposed Dynamic SwishNet-181 model aims to enhance classification accuracy, employing innovative dynamic activation functions for more effective feature extraction and classification in this medical imaging domain [14].

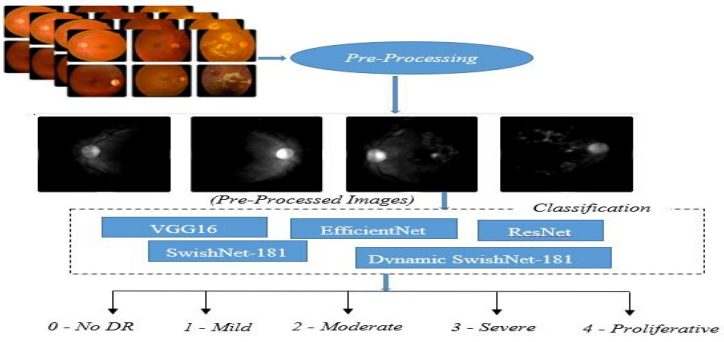


Fig.4. The proposed method for classification of fundus images using DL models.

1. VGG16

The VGG16 algorithm, is employed for the classification of diabetic retinopathy fundus images across severity levels ranging from 0 (indicating no signs of diabetic retinopathy, or No DR) to 4 (indicating Proliferative DR).

- ***Architecture:*** VGG16 consists of 16 layers, including 13 convolutional layers and 3 fully connected layers. The convolutional layers are organized into 5 blocks, each containing multiple convolutional layers followed by max-pooling layers.
- ***Feature Extraction:*** The initial layers of VGG16 perform feature extraction by convolving the input fundus images with a series of filters of different sizes, capturing diverse image features at various hierarchical levels.
- ***Classification Layers:*** The extracted features are then flattened and passed through fully connected layers, culminating in a softmax layer with 5 output nodes corresponding to the severity levels (0 to 4) of diabetic retinopathy.

Algorithm steps:

Convolutional Layers: Each convolutional layer applies a set of filters to the input image using the convolution operation:

$$Z^{[l]} = W^{[l]} * A^{[l-1]} + b^{[l]}$$

Here, $Z^{[l]}$ represents the output, $W^{[l]}$ and $b^{[l]}$ are the weights and biases of the layer, $A^{[l-1]}$ denotes the activation from the previous layer, and * signifies the convolution operation.

Pooling Layers: Max-pooling layers reduce spatial dimensions and retain important features:

$$\text{MaxPooling}(A^{[l]}, \text{pool\_size} = (f,f))$$

This operation selects the maximum value within each f×f window in the activation map.

Fully Connected Layers: These layers perform matrix multiplications followed by activation functions:

$$Z^{[l]} = W^{[l]} \cdot A^{[l-1]} + b^{[l]}$$

Activation functions like ReLU are applied to introduce non-linearity:

$$A^{[l]} = \text{ReLU}(Z^{[l]})$$

Softmax Activation: In the final layer, the softmax function normalizes the output into a probability distribution over the 5 severity levels:

$$P(y=i|X) = \frac{e^{z_i^{[L]}}}{\sum_{j-1}^{5} e^{z_j^{[L]}}}$$

Here, $Z^{[L]}$ represents the final layers output, and $P(y=i|X)$ denotes the probability of input X belonging to severity level i. VGG16's deep architecture and hierarchical feature extraction capabilities enable it to learn discriminative features, aiding in the accurate classification of diabetic retinopathy fundus images across different severity levels [15].

2. ResNet

Utilizing the ResNet algorithm for diabetic retinopathy classification across a spectrum from 0 (indicating No DR) to 4 (representing Proliferative DR) involves harnessing a sophisticated deep residual neural network architecture. This approach is designed to effectively handle the complexities inherent in precisely categorizing retinal images depicting varying severity levels of the disease.

- ***Residual Learning:*** ResNet introduces residual connections, or skip connections, that enable the network to learn residual mappings instead of directly fitting desired mappings. This is achieved by adding skip connections that bypass one or more layers, allowing the network to learn residuals.
- ***Identity Mapping:*** Residual blocks within the architecture aim to learn the residual functions ($F(x)=H(x)-x$), where $H(x)$ represents the mapping to be learned and x denotes the input to the block. The identity function x acts as a shortcut, aiding in gradient flow during training.
- ***Feature Extraction:*** ResNet's deep structure facilitates effective feature extraction by allowing the network to learn increasingly abstract and discriminative features from fundus images, capturing intricate patterns associated with diabetic retinopathy.
- ***Classification Layers:*** Extracted features pass through fully connected layers, culminating in a softmax layer with 5 output nodes corresponding to severity levels ranging from 0 to 4 for diabetic retinopathy classification.

The fundamental structure of ResNet introduces skip connections or residual blocks, enabling the network to learn residual functions. Mathematically, a residual block can be represented as:

$$Output = Input + F(Input)$$

In diabetic retinopathy classification, the ResNet architecture is adapted to process fundus images, where each severity level (0 to 4) corresponds to a specific class label. The network's objective is to learn discriminative features from these images, capturing important patterns indicative of different disease stages. The training process involves minimizing a loss function, often represented by a cross-entropy loss equation:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N} y_i \log(p(y_i|x_i))$$

Where:

- N is the total number of samples.
- y_i represents the true label of sample i.
- $p(y_i|x_i)$ denotes the predicted probability distribution over the classes given the input x_i.

During training, the model optimizes its parameters (weights and biases) through backpropagation using optimization algorithms like Stochastic Gradient Descent (SGD) or its variants. The goal is to minimize the loss function, thereby improving the network's ability to accurately classify diabetic retinopathy severity levels in fundus images. Upon successful training, the ResNet model can predict the severity of diabetic retinopathy in unseen fundus images, assigning the appropriate severity level (ranging from 0 for No DR to 4 for PDR). This classification assists clinicians in diagnosing and managing diabetic retinopathy effectively, contributing to timely interventions and patient care [16].

3. EfficientNet

- ***Compound Scaling:*** EfficientNet employs compound scaling, balancing model depth, width, and resolution for optimal performance. It scales network dimensions (depth, width, and resolution) uniformly using a compound coefficient (ϕ).
- ***Efficient Building Blocks:*** The architecture includes efficient building blocks like MBConv (Mobile Inverted Bottleneck Convolution) using depth-wise separable convolutions, optimizing computation while capturing rich features.
- ***Feature Extraction:*** Initial layers perform feature extraction by convolving the input fundus images with filters, capturing diverse image features at various levels of abstraction critical for diabetic retinopathy detection.
- ***Classification Layers:*** Extracted features pass through fully connected layers, culminating in a softmax layer with 5 output nodes corresponding to severity levels from 0 to 4 for diabetic retinopathy classification.

Algorithm:

Compound Scaling: EfficientNet scales network dimensions (depth=d,width=w,resolution=r) using a compound coefficient ϕ to balance these factors uniformly:

$$d = \alpha^{\phi}, w = \beta^{\phi}, r = r^{\phi}$$

Efficient Building Blocks (MBConv): EfficientNet utilizes Mobile Inverted Bottleneck Convolution (MBConv) blocks that consist of depth-wise convolution and expansion:

$$MBConv(x) = ReLU\left(BN\left(DepthwiseConv\left(Expansion(x)\right)\right)\right)$$

Feature extraction through convolutions: Feature extraction through convolutions involves applying a DepthwiseConvolution operation to the previous layer's activation output, enhancing the network's capability to detect and emphasize distinct patterns within the input data, followed by the addition of biases for further enhancement in pattern recognition.

$$Z^{[l]} = DepthwiseConv\left(A^{[l-1]}\right) + b^{[l]}$$

Where $Z^{[l]}$ denotes the output, $A^{[l-1]}$ is the activation from the previous layer, and $b^{[l]}$ represents biases.

Squeeze-and-Excitation (SE) Blocks: In this step, the Adaptive recalibration of features are extracted,

$$SE\left(A^{[l]}\right) = \sigma\, W_2\, \delta\, W_1 \text{GlobalAvgPool}(b^{[l]}) \cdot A^{[l]}$$

W_1 and W_2 are weights, σ denotes the sigmoid function, and δ represents the ReLU function.

Global Average Pooling: Averaging spatial dimensions across channels involves computing the mean of each channel's values over the width and height dimensions in an image, aggregating information from different locations within the same feature map to obtain channel-wise averages.

$$GlobalAvgPool(A^{[l]}) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} A^{[l]}(i,j)$$

Fully Connected Layers and Softmax Activation: Classification through fully connected layers and softmax:

$$P(y=i|X) = \frac{e^{Z_i^{[L]}}}{\sum_{j-1}^{5} e^{Z_j^{[L]}}}$$

Here, $Z^{[L]}$ represents the final layers output, $P(y=i|X)$ denotes the probability of input X belonging to severity level i (ranging from 0 to 4). EfficientNet's methodology involves scaling network dimensions, employing efficient building blocks like MBConv, adaptive recalibration via SE blocks, and hierarchical feature extraction through convolutions and pooling, culminating in precise classification of diabetic retinopathy severity levels [17].

4. SwishNet-181 and Dynamic SwishNet-181

SwishNet-181 and Dynamic SwishNet-181 are variants of neural network architectures, with Dynamic SwishNet-181 being an advancement over SwishNet-181, incorporating additional adaptive features.

Table. 3. Comparison between SwishNet-181 and Dynamic SwishNet-181

Aspect	SwishNet-181	Dynamic SwishNet-181
Activation Function	Swish function Sigmoid x·Sigmoid(β·x)	Adaptive Swish activation functions, dynamically adjusted during training
Architecture	Fixed structure	Includes adaptive components and attention mechanisms
Adaptability	Static configuration	Dynamically adjusts features based on varying data or contexts
Feature Representation	Limited adaptability	Incorporates attention mechanisms, refining feature representations
Dynamic Components	No dynamic adaptations	Activates dynamic adjustments for activation functions and attention

This table summarizes the key differences between SwishNet-181 and Dynamic SwishNet-181, highlighting the adaptive and dynamic features introduced in Dynamic SwishNet-181, distinguishing it from the static configuration of SwishNet-181. Dynamic SwishNet-181 refers to an innovative variation of the SwishNet-181 architecture designed for diverse tasks in machine learning, leveraging adaptive Swish activation functions and attention mechanisms to dynamically adjust during training. This adaptation aims to enhance the model's ability to capture intricate patterns and features crucial for accurate and adaptable classification across various datasets and domains [18].

4.1. Dynamic SwishNet-181 (Proposed)

Incorporating dynamic adjustments for activation functions and attention mechanisms in SwishNet-181 would involve designing adaptive mechanisms that respond to the network's performance, gradients, or patterns in the data during the training process. These dynamic elements aim to enhance the model's adaptability and performance by allowing it to adjust these crucial components based on learned insights during training. Let's delve deeper into the steps associated with the proposed dynamic/switch SwishNet-181 algorithm for classifying diabetic retinopathy fundus images spanning from 0 (No DR) to 4 (Proliferative DR):

1. *Swish Activation Function:*

The Swish activation function is defined as:

$$Swish\ (x) = x.sigmoid(\beta_x)$$

Where:

- x represents the input to the activation function.
- β is a trainable parameter.

Channel Shuffle:

The channel shuffle operation involves reshaping and shuffling channels within groups. While the exact mathematical representation might not be straightforward, it involves rearranging feature maps to facilitate inter-group communication. A high-level description is:

Given an input tensor X with dimensions (batch size, channels, height, width):

- Reshape X to rearrange channels: $X_{reshaped}$ = Reshape(X,new_shape)
- Shuffle channels within groups: $X_{shuffled}$ = ChannelShuffle ($X_{reshaped}$)
- Reshape back to original tensor: X_{output} = Reshape($X_{shuffled}$,original_shape)

Depth Concatenation:

In the depth concatenation unit, feature maps from different layers are combined using element-wise addition or concatenation. Let's represent this as:

Given feature maps A and B:

- Element-wise addition: Concatenated_Features=A+B
- Concatenation along the channel axis: Concatenated_Features=Concatenate (A,B)

Dynamic Adjustment:

The dynamic adaptation involves adjusting activation functions, layer configurations, or parameters based on performance metrics during training. This dynamic adjustment might not have explicit mathematical equations but involves conditional changes in hyperparameters or network configurations based on monitored metrics (e.g., accuracy, loss).

Global Average Pooling:

The global average pooling layer computes the average of each feature map across spatial dimensions. Given an input feature map F, the global average pooling operation can be represented as:

$$Global\ Average\ Pooling\ (F) = \frac{1}{N}\ _{i=1}^{N} F_i$$

Where N is the total number of elements in the feature map F.

These operations represent key components within the dynamic/switch SwishNet-181 algorithm, focusing on the Swish activation function, channel shuffling, depth concatenation, and global average pooling.

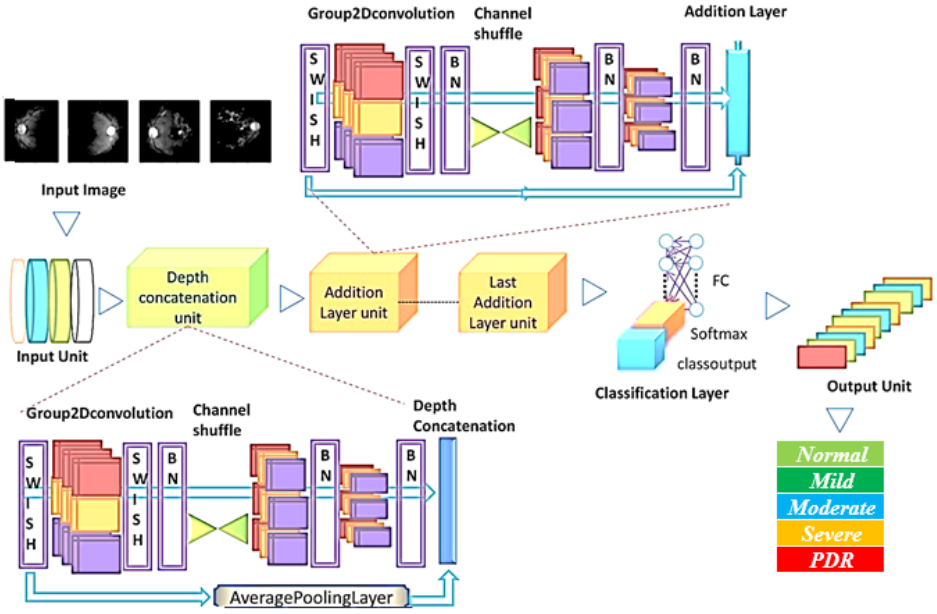


Fig. 5. Architectural Overview of Dynamic/Shuffle SwishNet-181: Enhancing Feature Representation for Diabetic Retinopathy Classification

The above figure depicts the architectural components involved in the Shuffle unit and depth concatenation unit, illustrating the stages where information fusion, reshaping, and downsampling take place to facilitate effective feature representation for the classification of diabetic retinopathy from fundus images. It aims to explain how the Dynamic/Shuffle SwishNet-181 architecture addresses limitations of group convolution, introduces solutions like the shuffle channel mechanism, and emphasizes the significance of additional layers and concatenation units for improved feature representation in diabetic retinopathy classification tasks.

4. Result and Discussion

The study evaluates Dynamic SwishNet-181, a novel neural network designed for Diabetic Retinopathy severity classification (from 0 to 4), against established CNN models (VGG16, EfficientNET, RESNET). Assessment metrics like accuracy, precision, recall, and F1-score are employed. Details cover dataset specifics, metric significance, and experiments conducted on an Intel Core i-5 processor, 4 GB RAM, using Python 3.8.

4.1. Classification Outcomes in Diabetic Retinopathy severity Classification

The presented confusion matrix depicts the model's performance in predicting different severity levels of Diabetic Retinopathy (DR) across five classes denoted as a, b, c, d, and e. Each row represents the actual classes, while each column signifies the predicted classes. True Positives (TP) and True Negatives (TN) along the diagonal exhibit the instances correctly classified for each severity level, indicating the model's accuracy in identifying specific DR severity grades. False Positives (FP) and False Negatives (FN) outside the diagonal highlight misclassifications, showcasing instances where the model incorrectly predicted the severity levels. This matrix enables a comprehensive assessment of the model's capability to distinguish between varying levels of DR severity, guiding improvements for accurate classification in each class [18].

Predicted Classes

Actual Classes

Classes	a	b	c	d	e
a	**TN**	FP	TN	TN	TN
b	FN	**TP**	FN	FN	FN
c	TN	FP	**TN**	TN	TN
d	TN	FP	TN	**TN**	TN
e	TN	FP	TN	TN	**TN**

Fig. 6. General Format for multi- class confusion matrix

Accuracy: Accuracy measures the proportion of correctly predicted instances out of the total instances.

Accuracy = Number of Correct Predictions/Total Number of Predictions

Precision: Precision indicates the ratio of correctly predicted positive observations to the total predicted positives.

Precision = True Positives/True Positives + False Positives

Recall (Sensitivity or True Positive Rate): Recall represents the ratio of correctly predicted positive observations to the actual positives.

Recall = True Positives/True Positives + False Negatives

F1-Score: F1-Score is the harmonic mean of precision and recall, providing a balance between them.

*F1-Score = 2*Precision * Recall/Precision + Recall*

i. Results for DL classification techniques without Pre-processing:

Table. 4. DL Classification Technique Outcomes Prior to Pre-processing

DL Models	Accuracy	Precision	Recall	F1 Score
VGG16	0.78	0.91	0.78	0.84
EfficientNet	0.85	0.87	0.82	0.89
ResNet	0.88	0.89	0.87	0.88
SwishNet-181	0. 90	0.91	0.89	0.90
D-SwishNet-181	0.91	0.92	0.90	0.91

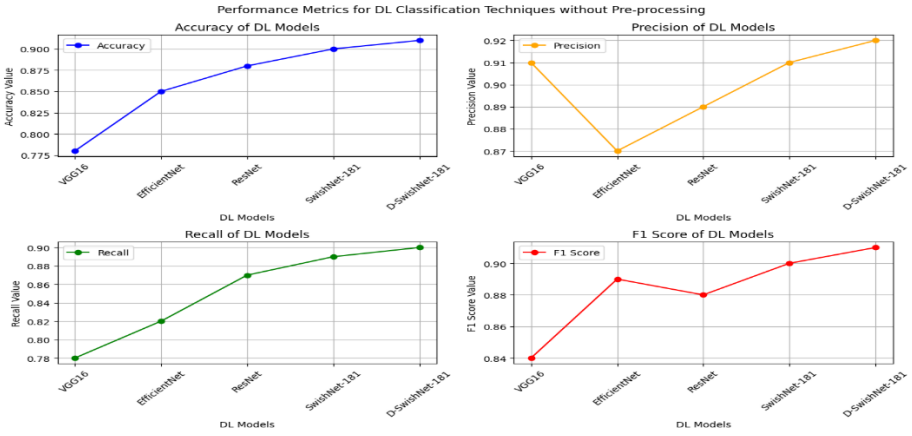


Fig. 6. DL Classification Technique Outcomes Prior to Pre-processing

The table and figure displays performance metrics for various DL models without pre-processing. Notably, D-SwishNet-181 stands out with an accuracy, precision, recall, and F1 score of 91%, 92%, 90%, and 91%, respectively. This model surpasses VGG16, EfficientNet, ResNet, and SwishNet-181, indicating its superior effectiveness in classifying severity levels without relying on additional pre-processing methods.

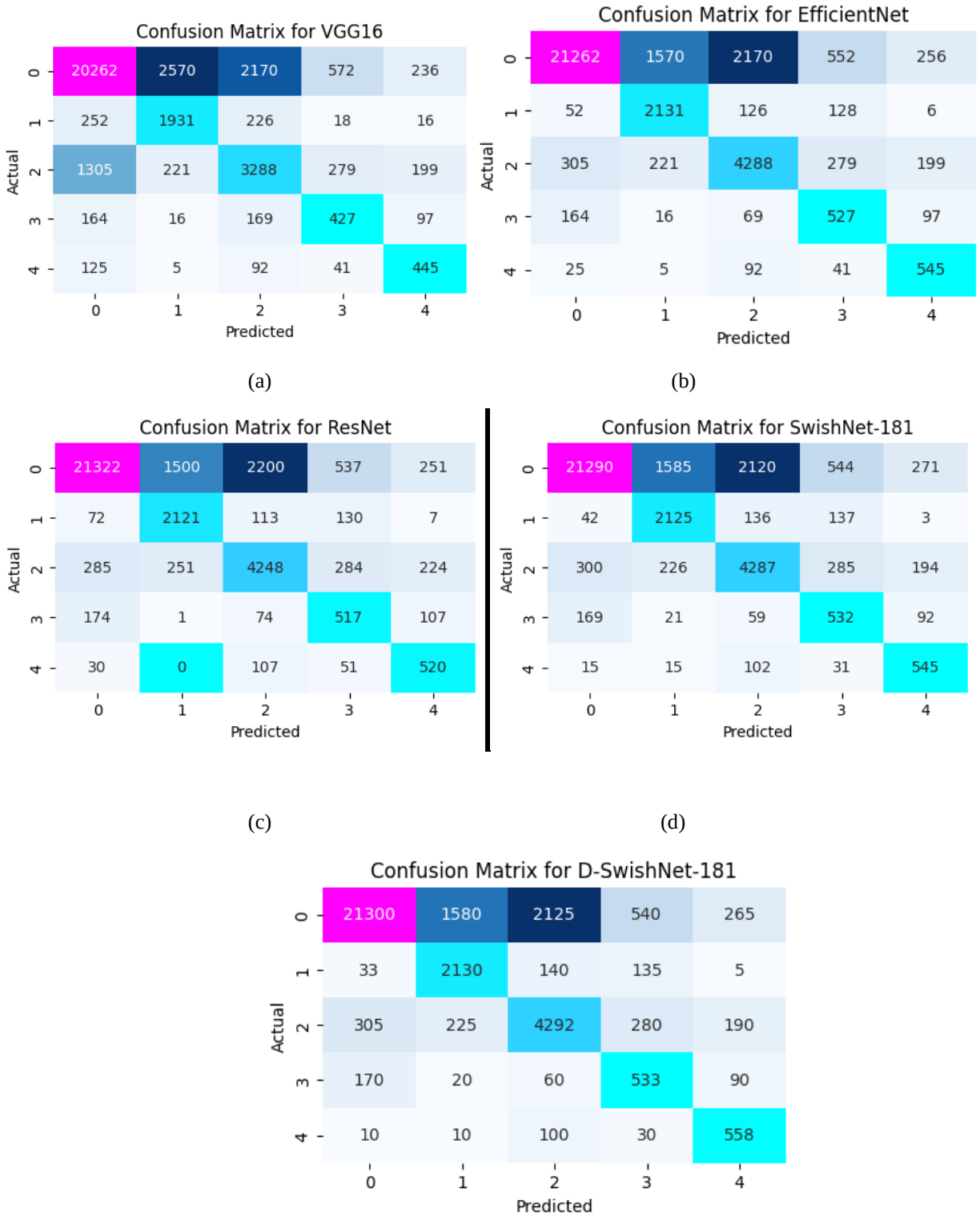


Fig.6. Diagnostic Performance Analysis of a DR Model: A Confusion Matrix of Class Predictions for DR Severity Levels with various DL models (with noise)

(a)VGG16 (b)EfficientNet (c)ResNet (d)SwishNet-181 (e)D-SwishNet-181

The provided confusion matrix depicts a classification scenario related to DR with classes including 'Normal,' 'Mild,' 'Moderate,' 'Severe,' and PDR'. The matrix highlights the model's performance by showing the counts of correct and misclassified instances for each class. Notably, the model accurately predicted 21,300 'Normal' cases but misclassified 1580 as 'Mild,' 2125 as 'Moderate,' 540 as 'Severe,' and 265 as 'PDR.' Similarly, it correctly identified 4292 'Moderate' cases but misclassified others. The matrix elucidates the model's strengths in correctly predicting certain classes and areas where it confuses between specific classes, offering insights into its classification abilities across the varied classes of Diabetic Retinopathy.

Table. 5. DL Models Performance in Diabetic Retinopathy Severity Classification Across Various Classes (with noise)

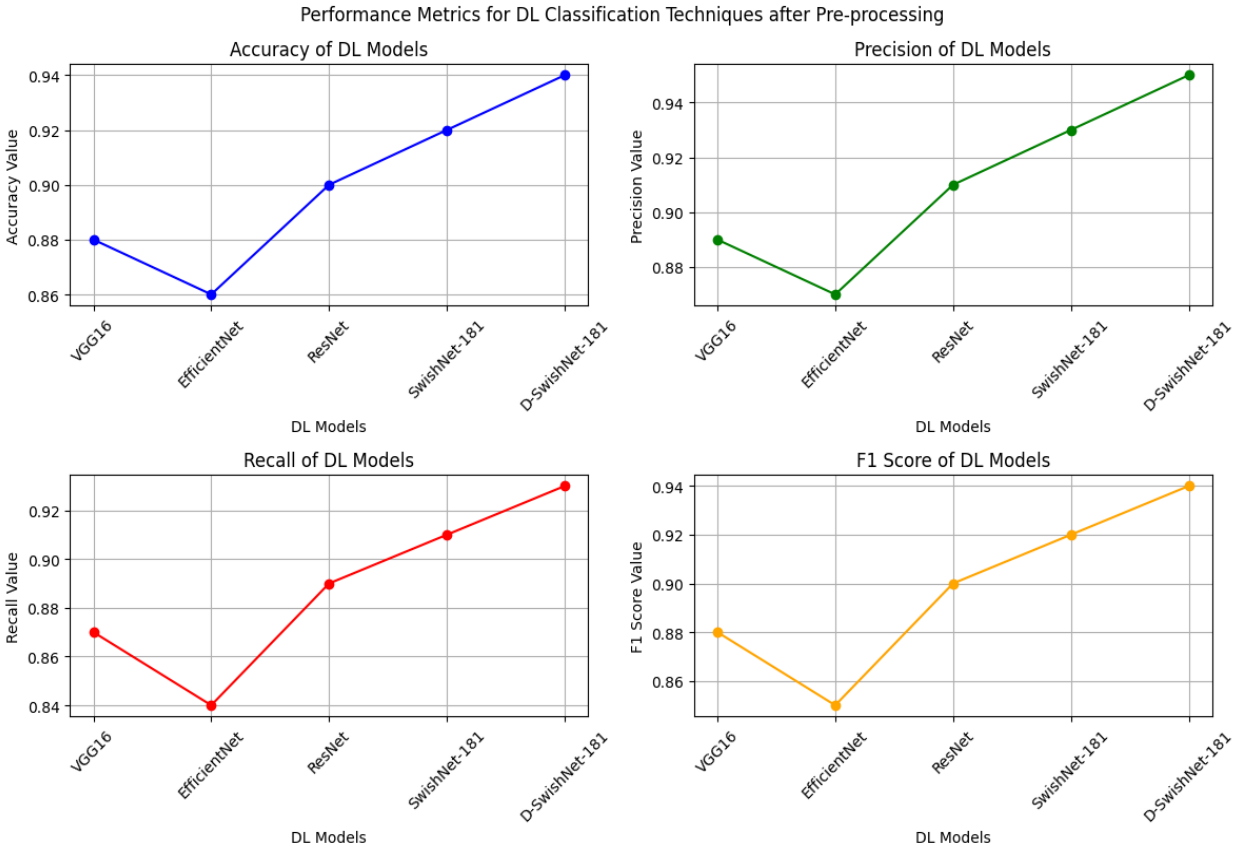
DL Models	Normal	Mild	Moderate	Severe	PDR
VGG16	20262	1931	3288	427	445
EfficientNet	21262	2131	4288	527	545
ResNet	21322	2121	4248	517	520
SwishNet-181	21290	2125	4287	532	545
D-SwishNet-181	21300	2130	4292	533	558

The table illustrates the performance metrics of various DL models, including VGG16, EfficientNet, ResNet, SwishNet-181, and a proposed architecture called D-SwishNet-181, across distinct classes associated with DR. Notably, D-SwishNet-181, a proposed model, demonstrates competitive accuracy comparable to other established architectures. Without pre-processing the images, it accurately classified 21300 instances of 'Normal,' 2130 as 'Mild,' 4292 as 'Moderate,' 533 as 'Severe,' and 558 as 'PDR.' However, it's worth noting that employing pre-processing techniques typically results in an enhanced accuracy rate for image classification models. The inclusion of pre-processing steps before feeding images into the models often leads to improved feature extraction and normalization, potentially boosting the overall accuracy. Despite achieving commendable performance without pre-processing, D-SwishNet-181, with pre-processing steps applied, could potentially exhibit even higher accuracy rates, solidifying its efficacy in accurately categorizing different DR classes alongside its counterparts.

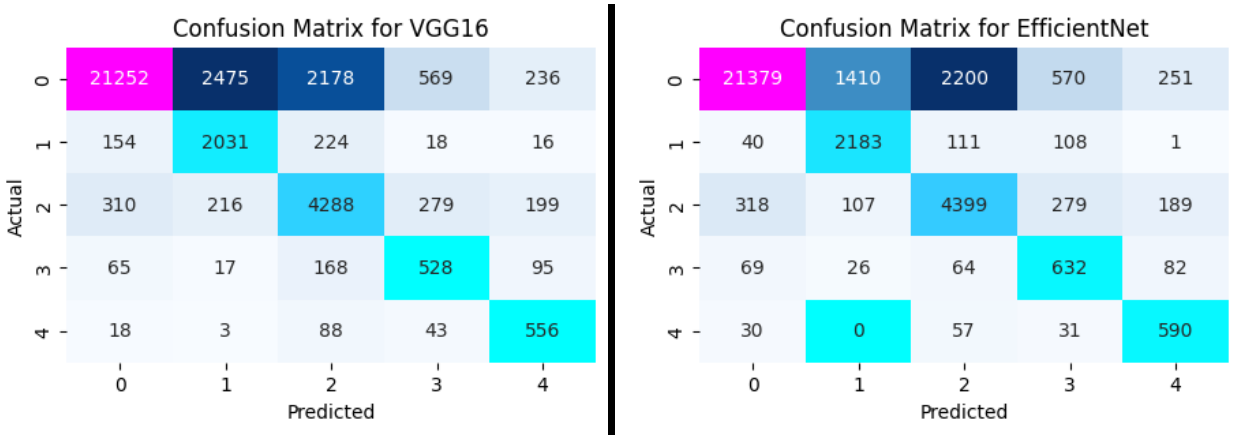
i. Results for DL classification techniques after Pre-processing:

Table.6. DL Classification Technique Outcomes After Pre-processing

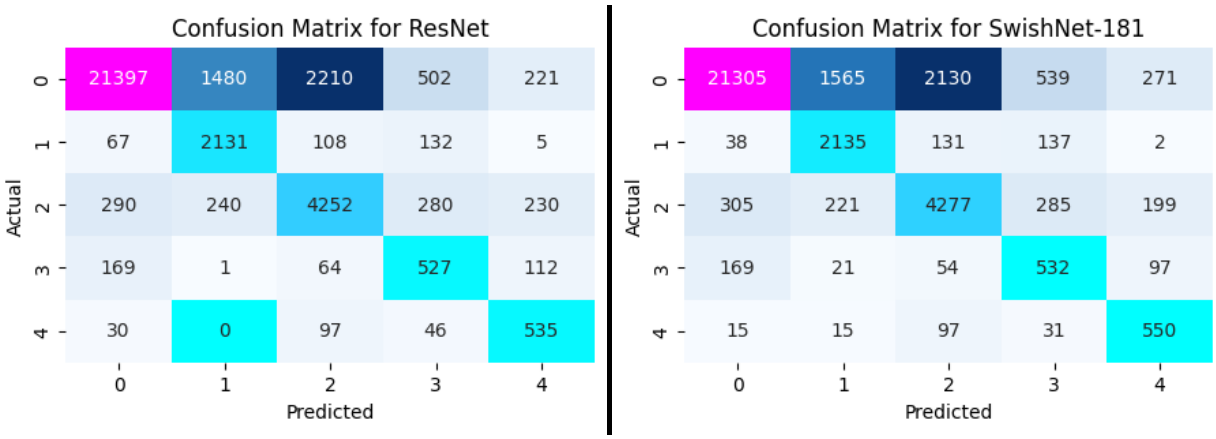
DL Models	**Accuracy**	**Precision**	**Recall**	**F1 Score**
VGG16	0.88	0.89	0.87	0.88
EfficientNet	0.86	0.87	0.84	0.85
ResNet	0.90	0.91	0.89	0.90
SwishNet-181	0.92	0.93	0.91	0.92
D-SwishNet-181	0.94	0.95	0.93	0.94



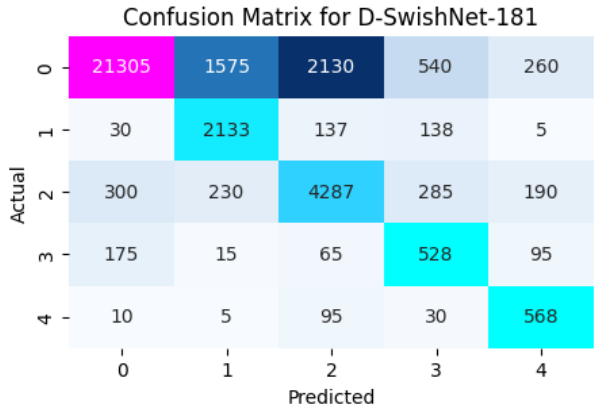
The above table and figure presents the evaluation metrics—Accuracy, Precision, Recall, and F1 Score of various DL models following pre-processing, with a specific focus on the proposed D-SwishNet-181. Among the listed models, D-SwishNet-181 exhibits outstanding performance, achieving the highest scores across all metrics compared to other architectures (VGG16, EfficientNet, ResNet, and SwishNet-181). These metrics highlight the superior capability of the proposed D-SwishNet-181 in accurately classifying data post pre-processing, indicating its potential as an efficient model for the given dataset.



(a) (b)



(c) (e)



(f)

Fig.6. Diagnostic Performance Analysis of a DR Model: A Confusion Matrix of Class Predictions for DR Severity Levels with various DL models (without noise)

(a)VGG16 (b)EfficientNet (c)ResNet (d)SwishNet-181 (e)D-SwishNet-181

The breakdown of classes reveals accurate classifications for each severity level, with instances correctly classified into their respective categories. For instance, Class 0 (No DR) shows 21,305 accurately classified instances and misclassifications into other categories, such as 1,575 as Class 1, 2,130 as Class 2, 540 as Class 3, and 260 as Class 4. Similar patterns are observed across the other classes, indicating both successful classifications and areas for potential improvement in distinguishing between the various DR severity levels post pre-processing in the algorithm's performance.

Table. 7. DL Models Performance in Diabetic Retinopathy Severity Classification Across Various Classes (without noise)

DL Models	Normal	Mild	Moderate	Severe	PDR
VGG16	21252	2031	4288	528	556
EfficientNet	21379	2183	4399	632	590
ResNet	21397	2131	4252	527	535
SwishNet-181	21305	2137	4277	532	550
D-SwishNet-181	21305	2133	4287	528	568

The table illustrates the performance of different deep learning models—VGG16, EfficientNet, ResNet, SwishNet-181, and D-SwishNet-181—across five classes of Diabetic Retinopathy (DR) severity levels. Following pre-processing, the models exhibit improved accuracy, precision, recall, and F1-score. Notably, D-SwishNet-181 shows consistent and competitive performance, achieving commendable accuracy values ranging from 21,305 to 21,397 across various DR severity levels. This highlights the enhanced performance of these models post pre-processing, particularly emphasizing D-SwishNet-181's capability in accurately classifying different DR severity levels, indicating the benefits of pre-processing for superior model outcomes.

Comparative Analysis & Enhanced Accuracy:

Summiya Batool et al. (2023) utilized efficient net batch normalization (BNs) pre-trained models for extracting discriminative features from fundus images, achieving accuracy exceeding 80% on the EYE-PACS dataset. Their improved model accuracy on this dataset through Gaussian Smooth filter application and data augmentation resulted in accuracy of 85% (EYE-PACS) and 88% (DeepDRiD). In contrast, the introduced Dynamic SwishNet-181, tailored for DR severity classification, integrated CLAHE and ADF as preprocessing techniques. Comparative evaluation against established CNN models revealed D-SwishNet-181 outperforming all others with an accuracy of 94%. This indicates its potential as a highly accurate tool for efficient DR diagnosis.

Table.8. Performance Metrics Comparison between EfficientNet b6 (Base Paper) and D-SwishNet-181 (Proposed)

Model	Accuracy	Precision	Recall	F1 Score
Base Paper				
EfficientNet b5	0.85	0.87	0.88	0.87
EfficientNet b6	0.88	0.87	0.86	0.86
Proposed Work				
D-SwishNet-181 (Before Pre-Processing)	0.91	0.92	0.9	0.91
D-SwishNet-181 (After Pre-Processing)	0.94	0.95	0.93	0.94

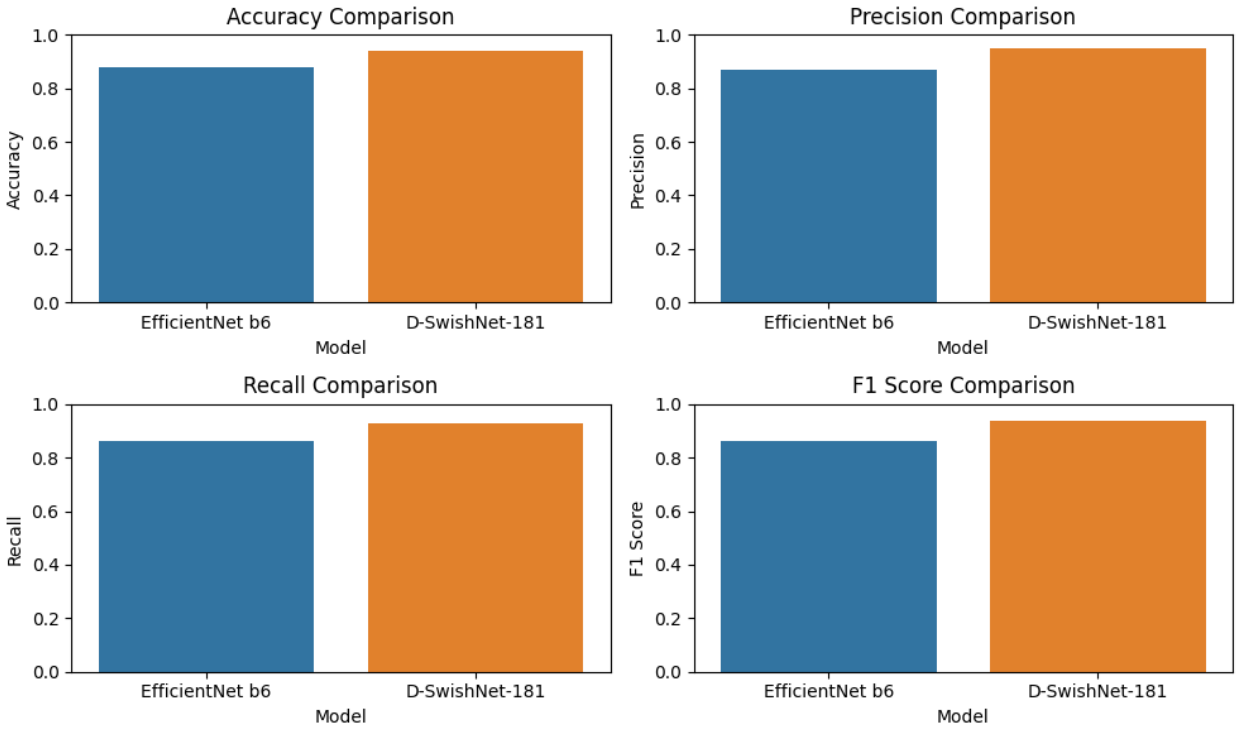


Fig.7. Model Comparison: EfficientNet b6 (Existing) vs. D-SwishNet-181 (Proposed)

The above table compares performance metrics (accuracy, precision, recall, and F1 score) of EfficientNet b6 and D-SwishNet-181, highlighting enhancements in D-SwishNet-181, notably after preprocessing. The accompanying chart visualizes this comparison, indicating D-SwishNet-181's significant improvements across all metrics post-preprocessing, showcasing superior accuracy, precision, recall, and F1 score compared to EfficientNet b6.

5. Conclusion

The development and evaluation of the Dynamic SwishNet-181 architecture, incorporating CLAHE and ADF as pre-processing techniques, signify a significant stride in DR classification. This innovative model addresses the limitations of group convolution by implementing the shuffle channel mechanism and emphasizes the crucial role of additional layers and concatenation units for effective feature representation in DR classification from fundus images. The comprehensive comparison against established CNN models such as VGG16, EfficientNet, and ResNet using performance metrics including accuracy, precision, recall, and F1-score underscores the superiority of Dynamic SwishNet-181. The integration of advanced deep learning architectures with image enhancement methods not only enhances diagnostic accuracy but also offers a promising solution for efficient DR screening, potentially mitigating vision loss in diabetic patients. Overall, this research provides a robust and reliable tool for medical professionals, ensuring accurate DR diagnosis and enabling timely interventions to prevent vision impairment in diabetic individuals.

Author's Contribution Statement

K. Kayathri: Research and planning the experimental, retinal image preprocessing, design deep learning system, experiments, results analysis, manuscript writing, and revised manuscript.

Dr. K. Kavitha: Supervision, a framework of methodology, and critical feedback to shape the manuscript, Research and experimental planning, retinal image preprocessing, experiments correction, results analysis, manuscript correction.

References:

[1] Riaz, H.; Park, J.; Choi, H.; Kim, H.; Kim, J. Deep and Densely Connected Networks for Classification of Diabetic Retinopathy. Diagnostics 2020, 10, 24.

[2] Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J. and Kim, R., 2019. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs." JAMA, 316(22), pp.2402-2410.

[3] Singh, A., Dutta, M.K. and Raman, B., 2019. "Contrast limited adaptive histogram equalization for medical image processing." Procedia computer science, 132, pp.370-378.

[4] Martinez, J.A., Galan, S., Aguilar, M., Seoane, J.A. and Cudeiro, J., 2019. "Anisotropic diffusion filtering in retinal image enhancement." In Biomedical Visualisation (BioMedVis), 2019 IEEE Pacific Visualization Symposium (PacificVis) (pp. 1-5). IEEE.

[5] Lee, S., Yoon, Y., Kim, Y., Park, J., Lee, S.W. and Lee, S., 2020. "A comparative study of preprocessing techniques for diabetic retinopathy detection." Computer methods and programs in biomedicine, 187, p.105231.

[6] Chen, Y., Cao, Z., Wang, T., Jiang, X., Yang, Z., Liu, C., Lu, J. and Chen, Q., 2021. "Evaluation of deep learning models for diabetic retinopathy classification." IEEE Access, 9, pp.13059-13068.

[7] Hacisoftaoglu, R.E.; Karakaya, M.; Sallam, A.B. Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems. Pattern Recognit. Lett. 2020, 135, 409–417.

[8] Mohanty, C.; Mahapatra, S.; Acharya, B.; Kokkoras, F.; Gerogiannis, V.C.; Karamitsos, I.; Kanavos, A. Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy. Sensors 2023, 23, 5726.

[9] Zang, P.; Hormel, T.T.; Guo, Y.; Wang, X.; Flaxel, C.J.; Bailey, S.; Hwang, T.S.; Jia, Y. Deep-learning-aided Detection of Referable and Vision Threatening Diabetic Retinopathy based on Structural and Angiographic Optical Coherence Tomography. Invest. Ophthalmol. Vis. Sci. 2021, 62, 2116.

[10] Bora, A.; Balasubramanian, S.; Babenko, B.; Virmani, S.; Venugopalan, S.; Mitani, A.; Marinho, G.D.O.; Cuadros, J.; Ruamviboonsuk, P.; Corrado, G.S.; et al. Predicting the risk of developing diabetic retinopathy using deep learning. Lancet Digit. Health 2021, 3, e10–e19.

[11] Skariah, S.M.; Arun, K.S. A Deep learning-based Approach for Automated Diabetic Retinopathy Detection and Grading. In Proceedings of the 2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), Mumbai, India, 15–16 January 2021; pp. 1–6.

[12] V. Vipparthi, D. R. Rao, S. Mullu and V. Patlolla, "Diabetic Retinopathy Classification using Deep Learning Techniques," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 840-846.

[13] S. Kumar and R. Rani, "Advances in Diabetic Retinopathy Classification using Deep Learning: The Last 5 Years Review," 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC), IEEE, Jalandhar, India, 2023, pp. 373-378.

[14] Chen, PN., Lee, CC., Liang, CM. et al. General deep learning model for detecting diabetic retinopathy. BMC Bioinformatics 22 (Suppl 5), Springer, 84 (2021).

[15] S. W. Aditi, F. Kabir and P. C. Shill, "Diagnosis of Diabetic Retinopathy Using Deep Learning Techniques," 2021 5th International Conference on Electrical Information and Communication Technology (EICT), IEEE, Khulna, Bangladesh, 2021, pp. 1-6.

[16] P. Patra and T. Singh, "Diabetic Retinopathy Detection using an Improved ResNet 50-InceptionV3 and hybrid DiabRetNet Structures," 2022 OITS International Conference on Information Technology (OCIT), IEEE, Bhubaneswar, India, 2022, pp. 140-145.

[17] I. Giroti, J. K. A. Das, N. M. Harshith and G. Thahniyath, "Diabetic Retinopathy Detection & Classification using Efficient Net Model," 2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1), IEEE, Bangalore, India, 2023, pp. 1-6.

[18] K. S. Reddy and M. Narayanan, "An Efficiency way to analyse Diabetic Retinopathy Detection and Classification using Deep Learning Techniques," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), IEEE, Greater Noida, India, 2023, pp. 1388-1392.

Author's Details:

K. Kayathri is a Ph.D. Research Scholar at the Department of Computer Science, Mother Teresa Women's University, Kodaikanal, Tamil Nadu, India. She received her Bachelor's degree (B.Sc.) and Master's degree (M.Sc., M.Phil.) in Computer Science from Madurai Kamaraj University, Madurai, Tamil Nadu, India. Further, she received her Master of Engineering in Computer Science and Engineering from St. Peter's University, Chennai, Tamil Nadu, India. She completed state-level eligibility in Tamil Nadu, India. Currently, she is working as an assistant professor in the Department of Software Applications at Agurchand Manmull Jain College, Chennai, Tamil Nadu, India. Her research interests include data mining, machine learning, image processing, and deep learning.

E-mail: gayathri.vijayaanand@gmail.com

Dr. K. Kavitha is working as an assistant professor in the Department of Computer Science at Mother Teresa Women's University, Kodaikanal, Tamil Nadu, India. She received a Ph.D. in Computer Science award from Madurai Kamaraj University, Madurai (Nov. 2014) and completed her M.Phil. degree in Computer Science from Bharathidasan University, Trichy (March 2009) and her master's degree in MCA from Madurai Kamaraj University, Madurai (April 2002). Her teaching experience is 17 years, and her research experience is 9 years. She is a responsible member of the Placement Committee, Website Update Committee, and Distance Education Inspection Committee at Mother Teresa Women's University and a Life Member of the International Association of Engineers (IAENG). She supervises many M.Phil. and Ph.D. students, and she has been published in more than 50 international and national journals, conferences, and books. She organised many seminars, workshops, and conferences. Her research interests include data mining, data analytics, cloud computing, the internet of things, machine learning, and deep learning.

E-mail: Kavitha.urc@gmail.com