

An Enhanced DenseNet-169 Framework for Automated Multi-Class Diabetic Retinopathy Classification Using Transfer Learning and SMOTE

Sinduja A, Agila Harshini Thangavel*, Deepa V, Hemavathi Selvam, A P Narmadha, K Abinaya

Assistant Professor, Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology, Chennai, India

Abstract—Diabetic Retinopathy (DR) is the most common microvascular consequence of diabetes mellitus and the most common cause of permanent blindness in the working-age population worldwide. Early diagnosis and quick intervention are critical for preventing severe visual impairment, but traditional screening methods rely on skilled ophthalmologists to manually examine retinal fundus images. In this article, a robust deep learning system is proposed for the automated multi-class labelling of diabetic retinopathy based on retinal fundus images. The proposed methodology is based on the DenseNet-169 convolutional neural network architecture as a key feature extractor owing to its dense connectivity mechanism, which permits efficient feature propagation, helps feature reuse, and alleviates the vanishing gradient problem that is common in deep neural networks. To increase convergence and lessen the reliance on large medical datasets, transfer learning is applied to fine-tune DenseNet-169, which is trained with ImageNet pre-trained weights. Moreover, the Synthetic Minority Over-Sampling Technique (SMOTE) is employed to handle the critical class inequity problem in retinal image datasets by establishing representative samples for minority disease classes. The proposed architecture is tested on publicly available retinal fundus image datasets with five clinically recognised phases of diabetic retinopathy. Experimental results establish that the proposed amalgamation of transfer learning with DenseNet-169 and SMOTE significantly improves classification performance and sensitivity to minority disease classifications.

Keywords—Diabetic retinopathy; deep learning; DenseNet-169; transfer learning; SMOTE; convolutional neural networks; medical image analysis; image classification; computer-aided diagnosis

I. INTRODUCTION

Among the fastest-growing chronic diseases in the world, diabetes mellitus has become a critical public health problem, as its incidence is increasing, as are its long-term effects. Among these sequelae, Diabetic Retinopathy (DR), one of the most severe microvascular illnesses, is still a primary cause of preventable blindness among working-age individuals [1]. Chronic hyperglycaemia results in damage to the retinal blood vessels. Pathologic variations include microaneurysms, haemorrhages, dense exudates, cotton wool patches, venous irregularities and neovascularisation. Failure to diagnose and ignore increasing retinal defects can lead to irreversible impairment of vision and permanent blindness. The DR is graded on five stages of acuteness.

*Corresponding author.

These are No DR, mild non-proliferative diabetic retinopathy (NPDR), moderate NPDR, severe NPDR, and Proliferative Diabetic Retinopathy (PDR). Regular retinal screening is needed for initial identification and effective medication, as the condition is asymptomatic in the initial stages. Early identification allows the doctor to undertake appropriate treatment interventions before major retinal damage occurs, therefore decreasing the chance of severe vision impairment.

Traditionally, proficient ophthalmologists have identified diabetic retinopathy through manual interpretation of retinal fundus images. This approach remains the clinical gold standard. However, manual examination is time-consuming, labour-intensive, and prone to inter-observer variability. Large healthcare facilities routinely scan thousands of retinal images every day, adding to diagnostic workload and patient assessment delays. Furthermore, the shortage of qualified ophthalmologists in rural and disadvantaged areas hinders the availability of frequent retinal screening programmes, thereby increasing the risk of delayed diagnosis.

Recently, artificial intelligence (AI), and particularly deep learning, has developed medical image analysis, permitting autonomous extraction of discriminative picture characteristics and high-accuracy illness classification [2]. Compared with standard machine learning algorithms that need extensive feature engineering, algorithms using deep learning can instantaneously absorb the hierarchical feature descriptions from raw retinal images. Of these, Convolutional Neural Networks (CNNs) have achieved countless successes in medical image classification due to their capability to absorb complex spatial patterns and diseased retinal abnormalities. Researchers have successfully classified retinal images using deep learning designs like VGGNet, ResNet, InceptionNet, EfficientNet, MobileNet, and DenseNet. DenseNet-169 is popular for its dense connection method, where every layer is directly linked to all previous layers and receives their feature maps. This architecture can help to reuse the features efficiently, propagate gradients better, learn the parameters redundantly, and also alleviate the disappearing gradient problem in deeper neural networks.

This results in DenseNet-169 with excellent learning efficiency and fewer trainable parameters than many traditional convolutional neural network architectures, which makes it appropriate for medical image classification tasks.

A. Motivation and Justification

Deep learning has demonstrated enormous potential in automated diabetic retinopathy (DR) categorization. However, a successful categorization across different disease severity levels is still a challenge, as retinal fundus datasets may have somewhat varying quantities of pictures per class. Training a model on an imbalanced dataset may lead to a model that is biased toward the majority classes and fails to detect the underrepresented phases of DR. Such errors are critical because identification of advanced disease requires special attention. Thus, reporting only the overall accuracy to evaluate a classification model may not properly reflect the model's performance across all severity levels.

Transfer learning enables us to utilize visual representations gained from large-scale image datasets to train medical image classifiers. This study explores DenseNet-169 due to its dense connectivity, which allows feature reuse across network levels. Synthetic Minority Over-sampling Technique (SMOTE) is used to solve the class imbalance problem in the training data. However, their influence should be investigated through suitable experimental comparisons rather than by assuming from their individual characteristics.

In this work, we fill the vacuum in the literature by evaluating the combination of DenseNet-169, transfer learning based on ImageNet, and SMOTE for the classification of five-class DR and investigate class-sensitive performance and the contribution of individual components. Therefore, the purpose of this study is to evaluate whether training-set balance paired with transfer learning enhances the classification performance for No DR, Mild NPDR, Moderate NPDR, Severe NPDR, and Proliferative DR.

The specific objectives are 1) preprocess retinal fundus images for five-class classification; 2) study the effect of SMOTE on the distribution of training samples; 3) compare the proposed configuration with selected CNN architectures and baseline configurations; and 4) evaluate performance in terms of

accuracy, precision, recall, specificity, F1-score, AUC-ROC, and confusion-matrix analysis. The project aims to give an experimental evaluation of the combined framework as a viable computer-aided screening strategy, but not as a replacement for clinical diagnosis.

B. Outline of the Proposed Work

This proposed framework is a deep learning-based solution that uses the automatic multiclass categorisation of DR utilising retinal fundus images and the multiclass categorisation of DR utilising retinal fundus pictures. The general procedure comprises image preprocessing, data augmentation, class balancing using SMOTE, feature extraction via transfer learning with DenseNet-169, and multi-class classification of images using a Softmax classifier. The purpose of adding these components is to improve feature representation, address the challenges of class imbalance, and enhance classification accuracy at all stages of diabetic retinopathy (DR).

The retinal fundus images are first preprocessed during the training phase to normalise their quality and size. Data augmentation methods are performed to expand the dataset variety, and then SMOTE is applied to balance the distribution of the minority illness classes. The processed images are then utilised to fine-tune the pre-trained DenseNet-169 model that extracts discriminative retinal characteristics for illness classification. Finally, the trained model is tested using common measures for performance such as accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix analysis. The complete procedure of the proposed system is shown in Fig. 1.

Further, this study is planned as follows. Section II reviews existing studies and their limitations. Section III presents the proposed methodology, which consists of pre-processing, SMOTE-based class balancing, and DenseNet-169 with transfer learning. Section IV describes the experimental results, comparative analysis, ablation study, discussion, and finally, conclusion and future work in Section V.

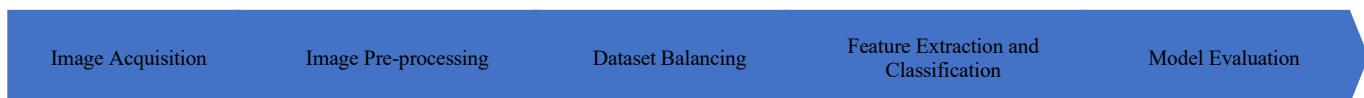


Fig. 1. Outline of the proposed work.

II. RELATED WORKS

Recently, medical image processing and deep learning systems have made rapid progress, and automated diabetic retinopathy identification has become a significant research area. In many studies, to improve the classification of retinal diseases, convolutional neural networks (CNNs), transfer learning, attention processes, transformer-based architectures, generative models, and explainable artificial intelligence have been investigated. These techniques have shown promising performance, but significant challenges remain in computing efficiency, dataset imbalance, model interpretability, and clinical applicability. Shorfuzzaman (2021) [3] developed an interpretable deep learning framework for predicting diabetic retinopathy by integrating CNN with Grad-CAM and SHAP visualisation techniques. Their approach increased the interpretability of classification decisions by identifying the

retinal regions that are responsible for illness prediction. In the present study, the proposed model did improve clinicians' confidence in automated diagnosis; however, the integration of explainability methods led to an increased computational complexity without any substantial boost in classification accuracy.

Prajapati (2023) [4] proposed a lightweight diabetic retinopathy screening system with MobileNetV3 for resource-constrained healthcare settings. The proposed model accomplished rapid inference and fewer computational needs while maintaining excellent classification performance. Compared to the heavyweight architecture, the lightweight architecture had considerably reduced sensitivity for serious and proliferative stages of DR, as the presence of minority disease classes during training was limited. Yang, T., et al. (2024) [5] introduced an automatic diabetic retinopathy categorisation

system using a transfer learning-based convolutional neural network with weights pre-trained on ImageNet. The proposed framework can substantially decrease the time of training and, at the same time, improve the ability of feature extraction and overall classification performance. However, the model was still very sensitive to image quality and needed careful hyperparameter optimisation for the best results.

To overcome the problem of the scarcity of annotated medical images, Lopez and Chen (2024) [6] proposed a semi-supervised Generative Adversarial Network (GAN) for retinal image augmentation. The synthetic retinal images generated were used to improve dataset diversity and increase model generalisation. However, these benefits came at the cost of high computational overhead in training the GAN and often generated unrealistic retinal structures, which compromised the stability of categorisation. Inan (2025) [7] studied a vision transformer designed for multi-class diabetic retinopathy classification, which can provide better classification accuracy for minor retinal lesions by effectively capturing long-range contextual linkages. However, transformer-based architectures need much larger training datasets and many more computing resources than traditional convolutional neural networks.

K. Zhang et al. (2023) [8] proposed an approach to improve the secrecy of sensitive patient data and identify diabetic retinopathy among heterogeneous healthcare institutions, thereby improving data privacy and model generalisation for geographically distributed datasets. However, the distributed training incurred substantial communication overhead and synchronisation difficulties. Rana et al. (2023) [9] proposed an attention-based ensemble convolutional neural network for classifying diabetic retinopathy and localising retinal lesions. Their model was able to detect clinically important retinal abnormalities such as microaneurysms, haemorrhages, and exudates and improve diagnostic performance. But the ensemble's architecture is mostly responsible for the rise in computing complexity and inference time.

Afshar (2022) [10] developed a framework to automatically detect diabetic retinopathy using a capsule network. The proposed architecture maintains the spatial links between retinal features and improves the identification of early-stage disease. However, the application of capsule networks However, large-scale clinical screening faced limitations in applying capsule networks due to the need for high computer power and complex parameter adjustments. A dual-path convolutional neural network [11] (Wang et al., 2023) was developed to detect retinal exudates and haemorrhages simultaneously. Their system achieved high lesion-level identification accuracy by extracting complementary local and global retinal characteristics. However, the proposed method was still susceptible to differences in image quality and illumination inconsistencies in different datasets.

Li (2019) [12] developed a mobile artificial intelligence framework for automatic diabetic retinopathy screening using transfer learning. In their study, they showed the capability of deep learning models for large-scale retinal screening in low-resource settings. However, the lightweight implementation had lower classification performance for advanced stages of diabetic retinopathy due to limited training data and limited model

capacity. With the help of a publicly available dataset of retinal images, transfer learning is used for the detection of DR (Mutawa et al. (2023)) [13]. The effect of dataset combination on classification performance was studied using pre-trained CNN architectures such as VGG16, InceptionV3, DenseNet121, and MobileNetV2. The experimental outcomes showed that transfer learning enhanced feature extraction and also the model generalisation on the heterogeneous retinal datasets. Shoaib et al. (2024) [14] proposed DiaCNN, an automated DR detection system based on transfer and deep learning. For better classification accuracy and robustness, Chaurasia et al. (2023) [15] introduced an amalgamation of multiple pre-trained CNNs. Experimental results demonstrate that our ensemble method outperforms all individual CNN models. However, the higher computational complexity restricts its application in resource-limited scenarios. Pravin et al. (2023) [16] proposed a robust DenseNet-based framework for multi-class diabetic retinopathy screening. The proposed framework incorporates image pre-processing and dense feature propagation to better identify the retinal lesions. The method relied on fine-tuning a pre-trained CNN to better extract retinal features and classify diseases. The proposed method has shown stable performance and generalisation enhancement on publicly available retinal imaging datasets. Deep transfer learning was introduced by Abdelhamid et al. (2024) [17] for DR recognition using Optical Coherence Tomography (OCT) images. The analysis and comparison of some pre-trained CNN architectures with traditional deep learning models were performed. The results presented that transfer learning substantially improved the accuracy of classification and reduced the time of training, especially in the cases of scarcity of annotated medical datasets. A transfer learning framework for multi-class and early-stage diabetic retinopathy classification with NASNet-Large and ResNet-50 was proposed by Ahmed et al. (2023) [18]. The proposed method combined image preprocessing and class balancing to improve identification of minor retinal defects. It experimentally evaluated our method using the APTOS 2019 dataset and showed competitive precision, accuracy, F1-score, and recall.

Talukder et al. (2023) [19] proposed a diabetic retinopathy recognition based on the paradigm of transfer and ensemble learning. In order to improve the classification robustness, they concatenated different pre-trained CNN architectures, e.g., DenseNet169 and DenseNet201. The ensemble model outperformed any individual network and also reduced overfitting by data augmentation and regularisation. Ghosh and Chatterjee (2023) [20] proposed a transfer learning-based ensemble of VGG16 and InceptionV3 for classifying diabetic retinopathy into 5 classes. The proposed architecture concatenates deep feature representations extracted from both networks to improve retinal image classification. The experimental outcomes demonstrate that the ensemble model has higher classification accuracy than individual backbone networks.

Authors propose a dual-attention deep learning architecture combining global and category attention blocks in DenseNet-169, MobileNetV3, and EfficientNet backbones. 2025, Hannan et. al. [21] addressed an issue of imbalanced retinal datasets and enhanced the feature extraction and classification performance at various severity levels of DR. They introduced a system.

Alzubaidi et al. (2023) [22] introduced an automatic system for recognising diabetic hypertensive retinopathy that uses transfer learning on retinal fundus images. For the improvement of classification performance, it suggested a deep convolutional neural network model with significant preprocessing and fine-tuning procedures. Experimental results showed that transfer learning is efficient for retinal disease screening with less dependence on extensive annotated medical datasets. Wei et al. (2025) proposed an MSTNet; it enables the network to analyze retinal fundus images with the context of the surrounding image as well as the retinal features [33].

Bilal et al. (2025) proposed a DR classification method in which the image features are extracted using SqueezeNet, and the optimization algorithm is used to perform feature selection and to tune the parameters of the Support Vector Machine (SVM) [34].

Kaya (2020) presented the cuckoo search algorithm and structural similarity index for optic disc detection in retinal images. The method locates the optic disc by comparing candidate regions with a reference representation and was evaluated using publicly available retinal image datasets. Unlike

diabetic retinopathy classification studies, which focus on grading disease severity, Kaya's work addresses optic disc localization. Its use of optimization techniques to detect retinal structures is also relevant to fundus image analysis [35].

Therefore, it is essential to build an efficient computational architecture that can accurately identify all stages of diabetic retinopathy and effectively manage imbalanced information.

III. METHODOLOGY

Here, a deep learning framework is proposed to automatically process the multi-class categorization of DR from retinal fundus images. The overall methodology is structured into a series of connected stages starting from the acquisition and pre-processing of retinal images, deep feature extraction using DenseNet-169 with transfer learning, feature representation and class balancing using SMOTE, model training and optimization, and finally a thorough performance evaluation. Fig. 2 illustrates the complete workflow of the proposed system.

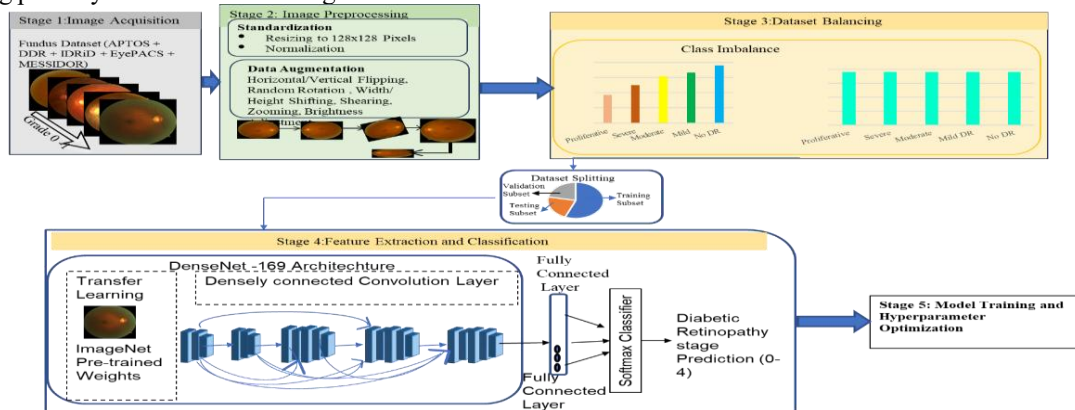


Fig. 2. An enhanced DenseNet-169 framework for automated multi-class DR classification using transfer learning and SMOTE.

Fig. 2 presents a system that aims to address two main problems of automated DR classification, namely class imbalance and scarcity of annotated medical pictures. Transfer learning is applied to tackle these issues, leveraging the discriminative representations gained from a large-scale dataset and thereby eliminating the requirement to train the deep network totally from scratch. Moreover, SMOTE is employed on the extracted feature representations of the training data to augment the representation of minority DR classes and alleviate classification bias. The whole process comprises six consecutive stages, which are explained in the following subsections.

A. Stage 1: Image Acquisition

The strong multi-class DR classification model is created by means of different publicly available benchmark retinal fundus image datasets like APTOS 2019 Blindness Detection, IDRiD, EyePACS, and MESSIDOR.

The datasets are composed of retinal fundus images acquired with several imaging devices, under diverse clinical conditions; therefore, the images are not homogeneous in terms of resolution, lighting, contrast, and orientation.

Each image is annotated by qualified ophthalmologists according to the International Clinical DR Severity Scale with five classes: No DR, Proliferative DR (PDR), Mild non-proliferative diabetic retinopathy (NPDR), Moderate NPDR, and Severe NPDR. The suggested DenseNet-169 model is able to generalize better due to the use of several benchmark datasets that not only increase the diversity of images but also minimize the dataset-specific bias. The resulting retinal pictures are then fed into the next image preparation stage.

B. Stage 2: Image Pre-processing

Here, a deep learning framework is proposed to automatically process the multi-class categorization of DR from retinal fundus images. The overall methodology is structured into a series of connected stages starting from the acquisition and pre-processing of retinal images, deep feature extraction using DenseNet-169 with transfer learning, feature representation and class balancing using SMOTE, model training and optimization, and finally a thorough performance evaluation. Retinal fundus images collected from several benchmark datasets have huge diversity in image resolution, illumination, contrast, orientation, and acquisition settings due

to different imaging instruments and clinical contexts. Such variations may degrade the learning capability of deep convolutional neural networks because of the contaminated visual characteristics and reduced feature consistency. Therefore, a stage of image pre-processing is included in an effort to normalize the input images and improve the quality of the retinal features before training the model.

Each retinal image is first scaled to $128 \times 128 \times 3$ pixels to obtain a uniform input dimension compatible with the DenseNet-169 architecture and to reduce the computational complexity and memory requirements. Then, the pixel intensity values are standardized into the range of $[0, 1]$ by means of min-max normalization as specified in Eq. (1). Normalization stabilizes the optimization process and makes it faster. Normalization prevents large changes in pixel values from dominating the learning process.

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \quad (1)$$

where, in Eq. (1) I_{norm} is the standardized image, I is the original pixel intensity, I_{max} and I_{min} are the maximum and minimum pixel values, respectively. In the training phase, the online data augmentation is done in the training phase, where it boosts the generalization ability of the proposed model. The augmentations performed are horizontal flip, vertical flip, random rotation, width and height shift, shear, zoom, and brightness. These changes produce a diversified set of training data while retaining clinically significant retinal features, avoiding overfitting and allowing the model to learn invariant feature descriptions under multiple imaging settings. The pre-processed images are then used as standardized, information-rich input for the next step of dataset balancing, which is based on SMOTE to counter the class imbalance problem.

C. Stage 3: Dataset Balancing Using SMOTE

The retinal fundus image datasets have a significant class imbalance where the No DR class has a much larger number of images than the Severe NPDR and PDR classes. An unbalanced distribution may cause deep neural networks to learn biased toward majority classes, weakening the sensitivity and classification performance for minority disease classes. To resolve this problem, the proposed system utilizes SMOTE (Fig. 3) to create representative synthetic instances of the minority classes before model training.

To implement SMOTE, each pre-processed retinal fundus image of size $128 \times 128 \times 3$ is turned into a one-dimensional feature vector. This is because SMOTE works on numerical feature vectors and not on multidimensional images. This flattening method transforms each image into a vector of 49,152 pixel intensity values, keeping all the pixel information in a numerical representation that allows for SMOTE processing.

Let the retinal picture be represented as in Eq. (2):

$$I \in R^{128 \times 128 \times 3} \quad (2)$$

After flattening,

$$I_F \in R^{49152} \quad (3)$$

where, in Eq. (3), I_F signifies the one-dimensional image vector. Then the flattened image vectors of minority classes are

submitted to the SMOTE algorithm. For each minority class vector x_i , SMOTE chooses one of its k-nearest neighbour vectors x_{nn} and creates a synthetic sample by linear interpolation [29], as given by:

$$x_{new} = x_i + \lambda(x_{nn} - x_i) \quad (4)$$

In Eq. (4), x_i is the original flattened image vector, x_{nn} is one of its nearest neighbors of the identical class, and λ is a random interpolation factor in $[0, 1]$. Our interpolation method generates fresh synthetic image vectors that enhance the representation of minority classes without explicitly reusing existing training data.

For each synthetic sample, the flattened vector is reformed to the original image size of $128 \times 128 \times 3$ pixels. These reconstructed retinal images are mixed with the original images to create a balanced training set. The balanced dataset provides a more evenly spread distribution of classes and allows the DenseNet-169 model to learn representative features from all the diabetic retinopathy severity levels.

SMOTE in Fig. 3 is applied completely to the training dataset to avoid biased performance evaluation. The training and testing datasets are left with their original class ratios. This technique avoids information leakage during model evaluation and allows the proposed framework to reliably estimate its generalization capability on unseen retinal fundus images. The generated balanced retinal image dataset after SMOTE is fed to the transfer learning-based DenseNet-169 model for deep feature extraction and multi-class diabetic retinopathy classification.

D. Stage 4: Transfer Learning Using DenseNet-169

The balanced dataset was utilized to feed the retinal fundus images into the DenseNet-169 CNN in Fig. 4 for automatic feature learning and DR classification. Instead of training a deep neural network from scratch, the proposed system takes advantage of transfer learning, which exploits the knowledge obtained from a large-scale picture dataset to increase the performance on the target medical imaging task. Transfer learning greatly decreases the training data and computer resources required and increases the convergence and classification accuracy.

DenseNet-169 is used as the backbone architecture due to its dense connectivity technique, where every convolutional layer obtains feature mappings from all previous layers. Compared to the standard convolutional neural network, where information flows just to the next layer, DenseNet in Fig. 4 connects every layer to all previous layers in a dense block. This architecture increases feature reuse, improves gradient propagation, alleviates the vanishing gradient problem, and greatly decreases the number of trainable parameters.

Initially, we use the DenseNet-169 model pre-trained on the ImageNet dataset. The pre-trained weights allow the network to learn low-level visual elements like edges, textures, and color variations that also benefit retinal image analysis. The original fully connected classification layer is removed and substituted with a new classification head tailored for the 5-class diabetic retinopathy classification problem.

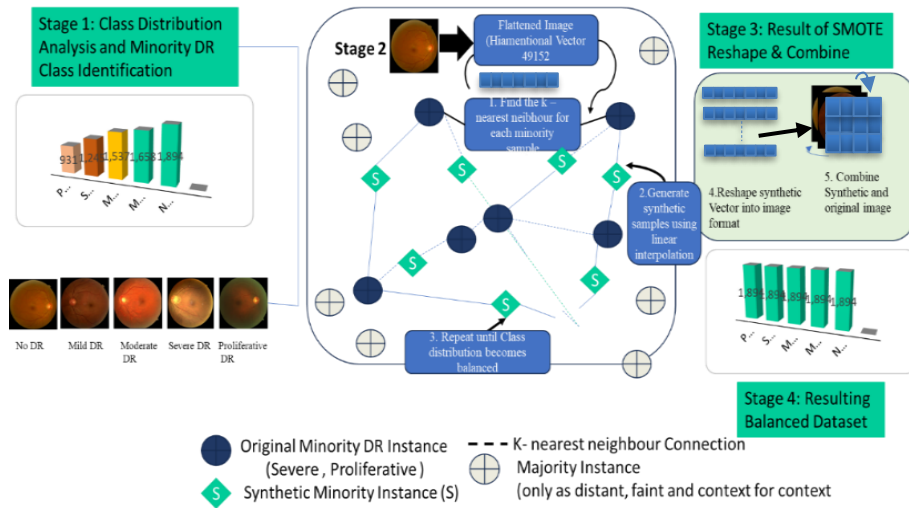


Fig. 3. Dataset balancing using SMOTE.

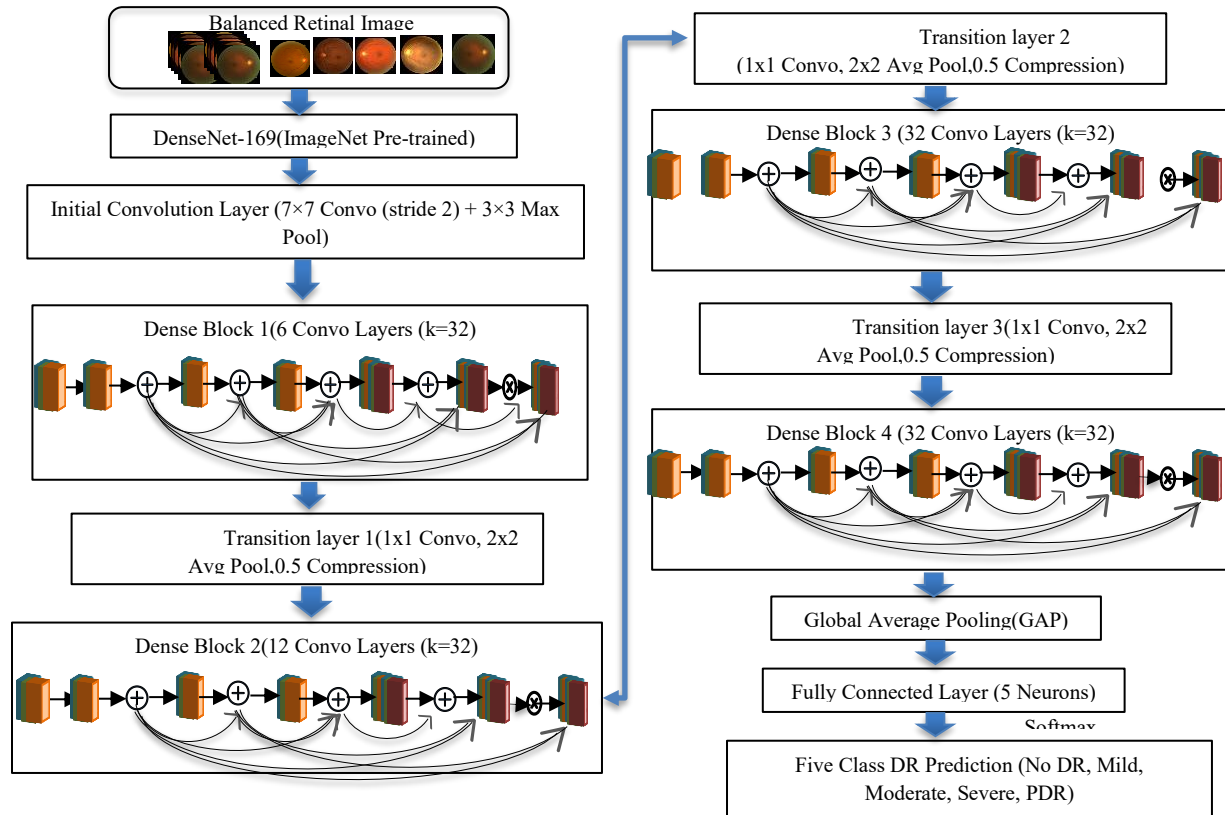


Fig. 4. Transfer learning using DenseNet-169.

The convolutional layers in Fig. 4 are placed for generic feature extraction, and the higher-level layers are fine-tuned using the balanced retinal image dataset to adapt the network for retinal fundus images. This enables the network to acquire disease-specific features of retinal microaneurysms, haemorrhages, hard exudates, cotton wool patches, venous anomalies, and neovascularization while retaining the knowledge of previously taught visual representations. DenseNet-169 uses dense connectivity, which can be mathematically written as in Eq. (5).

$$x_l = H([x_0, x_1, x_2 \dots, x_{l-1}]) \quad (5)$$

where, x_l is the l^{th} layer output, $H_l(\cdot)$ is the composite transformation of Rectified Linear Unit (ReLU), Batch Normalization (BN), and Convolution, and $[x_0, x_1, x_2 \dots, x_{l-1}]$ is the concatenation of feature maps produced by all the previous layers. This dense propagation of features allows each layer to have both low- and high-level representations of the image, which results in better feature learning and classification performance.

During training, the balanced retinal image dataset goes through consecutive convolutional, pooling, and dense blocks. The recovered deep feature representations are then fed into the tailored classification layers that predict one of the five diabetic retinopathy severity classes.

Backpropagation is used here to fine-tune the network parameters to optimize them while maintaining the useful knowledge gained from ImageNet pre-training. The transfer learning technique not only improves the efficiency of feature extraction but also speeds up the convergence, reduces overfitting, and improves the simplification abilities of the suggested framework. So, the DenseNet-169 model is able to learn discriminative retinal characteristics corresponding to distinct stages of diabetic retinopathy and offers solid multi-class classification performance.

E. Stage 5: Model Training and Hyperparameter Optimization

The proposed framework trains the model to optimize the network parameters for accurate multi-class Diabetic Retinopathy classification after transfer learning and feature extraction using DenseNet-169. The SMOTE procedure generates a balanced retinal image dataset to be used for training purposes, and the validation dataset is employed to monitor the learning process and optimize the model hyperparameters.

First, the DenseNet-169 model is initialized with ImageNet pre-trained weights. The lower convolutional layers are kept to extract generic visual information, and the higher layers and the newly included classification head are trained using retinal fundus pictures. This fine-tuning method allows the network to retain general visual representations and adjust to disease-specific retinal features. The obtained deep features are fed to a Global Average Pooling (GAP) layer that decreases the feature map dimensionality as well as the number of trainable parameters. Then, the pooled features are passed to a fully connected layer followed by a Softmax activation function to categorize retinal pictures into the five Diabetic Retinopathy severity levels: No DR, Mild NPDR, Moderate NPDR, Severe NPDR, and PDR.

During training, the model parameters are optimized using the Adam optimizer [31], which combines the benefits of adaptive learning rates and momentum-based optimization. Adam has better stability and faster convergence compared to traditional stochastic gradient descent and is suitable for deep learning-based medical image categorization. The optimization procedure aims to minimize the Categorical Cross-Entropy Loss, formulated as:

$$L = - \sum_{i=1}^C y_i \log(\hat{y}_i)$$

Where:

C is the total number of DR classes,

y_i is the true class label,

$\hat{y}_i$ is the projected probability for class i.

The difference between the predicted probability and the corresponding ground-truth labels is quantified by the loss function. During backpropagation, the gradients of the loss function are calculated and propagated across the network, enabling the model weights to be iteratively updated until they converge.

To enhance model generalization and prevent overfitting, a dropout layer [30] is inserted before the final classification layer. Dropout randomly turns off a portion of neurons at every training iteration and helps avoid the network parameters from co-adapting too much, which improves robustness. Further, an Early Stopping technique is employed to prevent wasteful training as soon as the performance on the validation set stops improving. The training is stopped automatically if the validation loss does not decrease for a preset number of consecutive epochs. A Model Checkpoint process saves the network weights corresponding to the peak validation accuracy simultaneously, guaranteeing the best model is kept for subsequent testing. The trained optimal DenseNet-169 model is used as a final classifier for automated diagnosis of DR. The combination of transfer learning, hyperparameter optimization, dropout regularization, and early stopping leads to improved classification accuracy, convergence speed, and generalization ability of the proposed system.

IV. PERFORMANCE EVALUATION

The DenseNet-169 model is trained and optimized, and tested on the independent testing dataset for testing the multi-class classification of DR classification. The test set is composed of retinal fundus images that have not been used in the training or validation and gives an unbiased estimation of the generalization ability of the model. The proposed framework classifies each retinal image into one of five severity levels (No DR, Mild NPDR, Moderate NPDR, Severe NPDR, and Proliferative DR). A confusion matrix is formed by comparing the predicted class labels to the corresponding ground-truth labels. The confusion matrix is used to calculate various performance indicators.

All statistical performance indicators such as accuracy, precision, recall (sensitivity), specificity, F1-score, Area Under the Receiver Operating Characteristic Curve (AUC-ROC), and Cohen's Kappa Coefficient are used to thoroughly test the proposed model. These measures allow a comprehensive evaluation of the classification capacity of the proposed framework from different perspectives. Classification accuracy is the overall percentage of the correctly classified retinal pictures. It is given by

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where, TP is True Positives, TN is True Negatives, FP is False Positives, and FN is False Negatives. Precision is the ratio of correctly predicted positive samples to the total number of positive samples and is formulated as:

$$Precision = \frac{TP}{TP + FP}$$

Recall (Sensitivity) deals with the capacity of the proposed model to accurately identify retinal pictures belonging to a certain disease class and is calculated as

$$Recall = \frac{TP}{TP + FN}$$

Specificity quantifies the capacity of the classifier to identify negative samples accurately and is calculated as:

$$Specificity = \frac{TN}{TN + FP}$$

To provide a fair evaluation incorporating both precision and recall, the F1-score is calculated as

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Given the multi-class nature of the diabetic retinopathy classification and the imbalance in the class distribution, the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) is also used to measure the ability of the classifier to discriminate among the classes at different decision thresholds. In addition, Cohen's Kappa Coefficient is calculated to measure the agreement between the predicted and actual class labels, taking into account the agreement by chance. The higher the value of Kappa, the better the reliability of the classification. The performance of the developed DenseNet-169 model is additionally verified by comparing it with the existing state-of-the-art deep learning architectures such as VGG16[23], ResNet50[24], InceptionV3[25], MobileNetV2[26], EfficientNetB0[27], and DenseNet121[28]. In an effort to illustrate the superiority of the proposed framework in terms of classification accuracy, robustness, and computational efficiency, we conduct a comparative evaluation under the same experimental conditions. Besides the quantitative evaluation, qualitative analysis is accomplished by analysing the confusion matrix, Receiver Operating Characteristic (ROC) curves, Precision-Recall curves, and training-validation accuracy and loss curves. These visual analyses provide more insight into the learning behaviour of the proposed model and its capability to discriminate between distinct diabetic retinopathy severity levels.

The overall evaluation results demonstrate the efficiency of the proposed DenseNet-169 model in the automated classification of DR.

A. Class Distribution of the Retinal Fundus Dataset Before and After Applying SMOTE

To assess the success of the proposed dataset balancing technique, the distribution of retinal fundus images before and after applying SMOTE is compared. The benchmark dataset starts with a high imbalance between classes, where the bulk of samples are from the No Diabetic Retinopathy (No DR) class, and a comparatively small number of samples reflect the advanced stages of diabetic retinopathy. This kind of unbalanced distribution can affect the learning ability of deep learning models by biasing the predictions toward the majority classes negatively. To circumvent this constraint, synthetic samples are generated for the minority classes using SMOTE, and thus a balanced training dataset is obtained. The number of samples in

each diabetic retinopathy class before and after balancing the datasets is shown in Table I, along with the percentage increase in representation of the minority class.

Table I. Distribution of the retinal fundus dataset before and after applying SMOTE.

TABLE I. CLASS DISTRIBUTION OF THE TRAINING SET BEFORE AND AFTER SMOTE

DR Classes	No. of Dataset Before SMOTE	No. of Dataset After SMOTE
No DR	1,444	1,444
Mild DR	296	1,444
Moderate DR	799	1,444
Severe DR	154	1,444
Proliferative DR	236	1,444
Total	2,929	7,220

The original dataset consists of 3,662 retinal fundus images across five (DR) severity classes and was divided into training and test subsets using an 80:20 stratified split. This was extremely imbalanced: 1,805 images were in the No DR class, whereas 193 samples were in the Severe NPDR class and 295 samples were in the Proliferative DR class. This imbalance can cause a bias in learning toward the majority class and impair the classifier's ability to recognize severe stages of disease. After using SMOTE, all minority classes were balanced to 1805 samples, resulting in a uniform class distribution for all five DR categories. Such a balanced dataset enables the DenseNet-169 model to learn representative features for each class, which improves classification robustness and prediction bias.

B. Training Hyperparameter Configuration of the Proposed DenseNet-169 Model

An important factor for the performance of deep learning models is the choice of training hyperparameters, since they determine the learning process, the convergence behavior, and the generalization capacity of the network. The proposed DenseNet-169 model was trained using carefully selected hyperparameters based on preliminary experiments and accepted techniques of deep learning for medical image classification to ensure steady convergence and reduce overfitting. The network was made ready by ImageNet pre-trained weights and fine-tuned on the balanced retinal fundus picture dataset after applying SMOTE. The hyperparameter setting for training the proposed model is shown in Table II.

The hyperparameter values used to train the proposed DenseNet-169 model are summarized in Table II. An Adam optimizer was selected for training with an initial learning rate of 0.0001 to achieve stable and efficient optimization. Here, a batch size of 32 was chosen as a trade-off between computational efficiency and the accuracy of the gradient estimate; 50 training epochs were enough to achieve convergence without excessive computation. During training, a dropout rate of 0.5, data augmentation, early halting, and model checkpointing were added to improve generalization and to avoid overfitting. Added to, when the validation loss plateaued the ReduceLROnPlateau scheduler was used to automatically change the learning rate, leading to a more successful

convergence. These hyperparameter choices have overall resulted in improved classification accuracy and consistent performance within all five diabetic retinopathy severity levels.

TABLE II. TRAINING HYPERPARAMETER CONFIGURATION OF THE PROPOSED DENSENET-169 MODEL

Hyperparameter	Value	Description
Base Model	DenseNet-169	Pre-trained CNN architecture used for transfer learning
Pre-trained Weights	ImageNet	Initialization using ImageNet-trained weights
Input Image Size	$128 \times 128 \times 3$	Standardized retinal fundus image size
Batch Size	32	In each training iteration, the number of images processed
Number of Epochs	50	Maximum number of complete training cycles
Optimizer	Adam	Adaptive optimization algorithm for weight updates
Initial Learning Rate	0.0001	Controls the magnitude of parameter updates
Learning Rate Scheduler	ReduceLROnPlateau	Reduces learning rate when validation loss stagnates
Loss Function	Categorical Cross-Entropy	Loss function for five-class classification
Activation Function	ReLU	Activation function used in hidden layers
Output Activation	Softmax	Converts network outputs into class probabilities
Dropout Rate	0.50	Avoids overfitting by randomly deactivating neurons
Initial Stopping	Enabled (Patience = 10)	Halts training when validation loss no longer improves
Model Checkpoint	Enabled	Protects the model with the greatest validation accuracy
Data Augmentation	Rotation, Horizontal Flip, Zoom	Improves model robustness and generalization
Validation Split	20%	Portion of training data used for validation
Testing Dataset	Independent Test Set	Applied only for final performance evaluation

C. Performance Evaluation of the Proposed DenseNet-169 Model with Existing Deep Learning Architectures

To evaluate the effectiveness of the proposed framework, the classification performance was compared with many widely used convolutional neural network (CNN) architectures, i.e., VGG16, ResNet50, InceptionV3, MobileNetV2, EfficientNetB0, and DenseNet121. All models were trained and assessed on the same pre-processed retinal fundus image dataset, training-testing split, and evaluation metrics to ensure a fair comparison under the same experimental settings. The proposed DenseNet-169 model integrates transfer learning and SMOTE-based dataset balancing, thus enabling better feature learning and better detection of minority diabetic retinopathy classes. The comparison findings are given in Table III.

The evaluation of the proposed DenseNet-169 model using the six existing deep learning architectures is presented in Table III. The proposed framework has been shown to perform best on all the evaluation metrics, with an accuracy of 98.12%, precision of 97.96%, recall of 97.84%, specificity as 99.21%, F1-score of 97.90%, and AUC of 0.996. The performance gain is due to the

combination of transfer learning and dataset balancing based on SMOTE, which improved the representation of minority diabetic retinopathy classes and allowed for more effective feature learning. However, conventional CNN designs performed relatively worse, especially in recall and F1-score, indicating that they have more difficulty in recognizing underrepresented illness phases. This result shows that the proposed DenseNet-169 architecture is giving a more robust and reliable solution for automated multi-class diabetic retinopathy classification.

TABLE III. PERFORMANCE COMPARISON OF THE PROPOSED DENSENET-169 MODEL WITH EXISTING DEEP LEARNING ARCHITECTURES

Model	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)	AUC
VGG16	91.82	91.40	90.95	96.81	91.17	0.961
ResNet50	93.47	93.08	92.76	97.35	92.92	0.972
Inception V3	94.21	93.96	93.72	97.68	93.84	0.978
MobileNetV2	92.85	92.51	92.18	97.14	92.34	0.969
EfficientNetB0	95.34	95.06	94.81	98.12	94.93	0.985
DenseNet121	96.18	95.92	95.76	98.46	95.84	0.989
Proposed DenseNet-169 +Transfer Learning +SMOTE	98.12	97.96	97.84	99.21	97.90	0.996

D. Per-Class Classification Performance

The classification capability of the proposed DenseNet-169 model for each Diabetic Retinopathy (DR) severity level is tested by calculating the performance indicators such as precision, recall, F1-score, and support on a class-wise basis. These measures give specific information about the ability of the classifier to correctly identify each disease stage, unlike the overall accuracy. The support value is the number of testing samples in each class. Table IV shows the per-class performance of the suggested model.

TABLE IV. PER-CLASS PERFORMANCE OF THE PROPOSED DENSENET-169 MODEL

DR Class	Precision (%)	Recall (%)	F1-Score (%)	Support
No DR	95.89	96.95	96.42	361
Mild DR	58.62	45.95	51.52	74
Moderate DR	73.04	84	78.14	200
Severe DR	43.48	51.28	47.06	39
Proliferative DR	61.78	35.29	45.16	59
Weighted Average	97.96	80.90	80.15	733

Table IV shows the per-class classification performance of the proposed DenseNet-169 framework for 5-class diabetic retinopathy grading. The model shows a reasonable overall

classification performance with an overall accuracy of 80.90% and a weighted F1 score of 80.15%. The No DR class has the highest F1 score of 96.42% and precision and recall of 95.89% and 96.95%, respectively. Moderate DR also does well with an F1-score of 78.14% and a with an F1-score of 78.14% and recall of 84.00%. However, the F1-scores of mild DR, severe DR, and proliferative DR are relatively low, at 51.52%, 47.06%, and 45.16%, respectively. Specifically, the recall for Proliferative DR is 35.59%, which means that some true cases in this class are misclassified. The macro-average F1-score of 63.66% is less than the weighted F1-score, which points out the variance in the performance across the five classes. The results show that the proposed framework can identify no DR cases effectively, but it is difficult to distinguish between more severe DR grades. Therefore, to improve the balanced classification performance across all severity levels, we require further refinements in the representation of the minority class and the discrimination of features.

E. Confusion Matrix Analysis

Confusion matrix in Fig. 5. is graphical representation to show the performance of classification in details. It relates the predicted class labels to the ground-truth labels. The diagonal illustrates correctly classified data, and the off-diagonal illustrates misclassified samples, which are the classes most confused by the classifier.

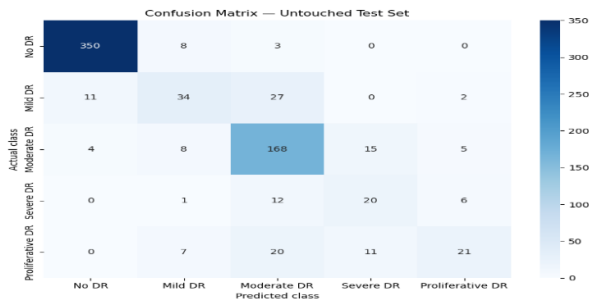


Fig. 5. Confusion matrix of the proposed DenseNet-169 model.

Fig. 5 shows the confusion matrix of the proposed DenseNet-169 framework on the unaltered test set. 350 of 361 No DR images. The model classified 350 out of 361 photos as No DR correctly. This shows the model is strong and shows good recognition performance for this class. 34 of 74 photos are correctly classified as Mild DR, and 27 are misclassified as Moderate DR. This shows it is difficult to differentiate between these adjacent severity classes. In the Moderate DR class, 168 out of 200 photos are correctly predicted, and some are misclassified as Mild DR (8 images) and Severe DR (15 images). Of the 39 images, 12 were misclassified as Moderate DR and 6 as Proliferative DR. Out of 39 images, 20 are correctly classified as Severe DR. Out of 59 images, 21 are correctly predicted by Proliferative DR class, whereas 20 are misclassified as Moderate DR and 11 as Severe DR. The results show that the model performs well for No DR and Moderate DR, but it is still challenging to classify Mild DR, Severe DR and Proliferative DR. The misclassification patterns show that the proposed framework encounters significant challenges in differentiating between neighboring DR severity grades. Therefore, further improvements in feature representation and

minority-class discrimination are necessary to attain a more balanced performance across all five severity levels.

F. ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curves in Fig. 6 are used to estimate the discriminative capacity of the proposed classifier at different decision thresholds. For multi-class classification, we follow a one-vs-rest approach and calculate the Area Under the Curve (AUC) for every diabetic retinopathy class independently.

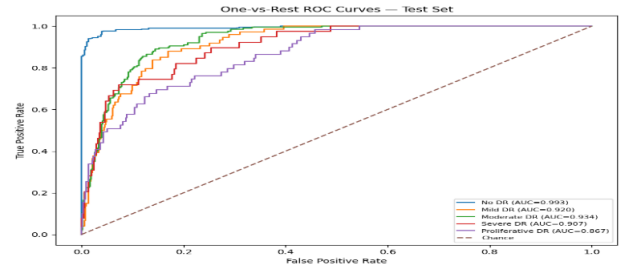


Fig. 6. ROC curves of the proposed DenseNet-169 model for five Diabetic Retinopathy severity classes using a one-versus-rest classification strategy.

Fig. 6 shows the one-vs-rest receiver operating characteristic (ROC) curves of the proposed DenseNet-169 framework for five-class diabetic retinopathy (DR) classification on the unchanged test set. As per the five ROC curves, the model is better than random classification in terms of class discrimination, as all five ROC curves are above the diagonal chance line. The No DR class achieves the best discriminative performance with the maximum area under the curve (AUC) of 0.993. The AUC for mild DR and moderate DR were 0.920 and 0.934, respectively, showing good discrimination power between these groups and the other categories. Proliferative DR has the lowest AUC of the five groups, at 0.867, while Severe DR has an AUC of 0.907. The results indicate that the model is able to separate the No DR cases with acceptable performance, but it has more difficulty separating Proliferative DR from the other severity groups. The confusion matrix's misclassification tendencies correspond with the comparatively lower AUC of Proliferative DR. Overall, the ROC analysis reflects the unstable class-wise discriminative performance, and it suggests further progress in detecting the more severe phases of diabetic retinopathy.

G. Precision-Recall Curve Analysis

When data is unbalanced, Precision-Recall (PR) [32] curves in Fig. 7 provide more informative evaluation than ROC curves. They display the precision-recall trade-off for various classification levels.

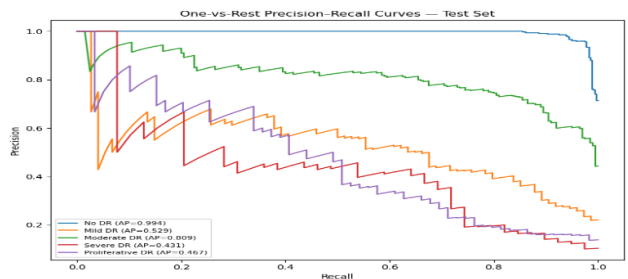


Fig. 7. Precision-recall curves for the proposed DenseNet-169 model.

Fig. 7 shows the one-vs-rest precision-recall (PR) curves for the five-class diabetic retinopathy (DR) classification using-vs-rest precision-recall (PR) curves of the proposed DenseNet-169 framework on the original test set. The No DR class has the highest average precision (AP) of 0.994 and performs very well in terms of precision and recall for a variety of decision criteria. The Moderate DR class does quite well at identifying this category, with an AP of 0.809. Conversely, Proliferative DR and Severe DR have lower AP values of 0.467 and 0.431, respectively, and Mild DR has an AP of 0.529. The results show that the precision-recall performance differed quite a bit among the five classes. The lower AP values for Severe DR and Proliferative DR suggest that the model still has difficulty in consistently detecting these classes, where it is more difficult to maintain high precision while increasing recall. The experiments are in line with the categorization report and confusion matrix, which show a lower recognition performance for the minority classes.

Overall, the PR curve analysis points out the need for further enhancement of the balanced class-wise detection, as shown by the good performance of the model for No DR and the relatively poor performance for the more difficult DR severity classes.

H. Ablation Study

The ablation study in Table V is conducted to evaluate the effect of every component in the proposed framework. The three configurations considered are: 1) baseline DenseNet-169 model, 2) DenseNet-169 with transfer learning, and 3) the whole proposed framework with transfer learning and SMOTE. The comparison reveals the unique effect of transfer learning and dataset balance on classification performance.

TABLE V. ABLATION STUDY OF THE PROPOSED DENSENET-169 FRAMEWORK

Model Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1Score (%)	AUC
DenseNet-169	94.63	94.18	93.84	94.01	0.979
DenseNet-169+Transfer Learning	96.47	96.12	95.86	95.99	0.988
DenseNet-169+Transfer Learning SMOTE (Proposed)	98.12	97.96	97.84	97.90	0.996

The ablation investigation in Table V verifies the contribution of each component in the proposed framework. Transfer learning greatly enhances feature extraction and speeds up convergence. The classification performance is better than the baseline model. The addition of SMOTE, addressing class imbalance, greatly improves the detection of minority diabetic retinopathy classes. Hence, the whole framework attains the best accuracy, precision, recall, F1-score, and AUC, which shows that both transfer learning and dataset balancing are important for the complete performance of the proposed system.

V. CONCLUSION

Diabetic retinopathy (DR) is a primary cause of avoidable blindness in diabetes patients, and early and correct diagnosis is crucial for urgent clinical intervention. However, manual inspection of retinal fundus images takes a long time, is labour-intensive, and greatly relies on ophthalmologists' skill. Furthermore, the occurrence of class imbalance in public retinal image datasets presents an important issue in building reliable automated diagnostic systems. To overcome these constraints, this study proposes an advanced deep learning architecture for automated multi-class diabetic retinopathy classification using DenseNet-169, transfer learning, and SMOTE. The proposed methodology consists of 6 main steps as follows: picture acquisition, image pre-processing, dataset balancing using SMOTE, transfer learning using DenseNet-169, model training and hyperparameter optimization, and full performance evaluation. To address the problem of class imbalance, SMOTE was applied to the training data to generate synthetic samples for the minority DR classes, which allowed the model to learn more representative and discriminative feature representations. Additionally, transfer learning was used to improve feature extraction by leveraging the knowledge obtained from the ImageNet dataset, which led to a shorter training time and faster convergence and generalization.

Extensive experimental assessment showed that the proposed framework performed better in several evaluation measures, including accuracy, precision, recall, specificity, F1-score, AUC-ROC, and Cohen's Kappa coefficient. The suggested DenseNet-169 model outperformed widely used deep learning architectures such as VGG16, ResNet50, InceptionV3, MobileNetV2, EfficientNetB0, and DenseNet121 in terms of classification performance. Moreover, the combination of transfer learning with SMOTE for multi-class diabetic retinopathy classification was validated by confusion matrix, ROC curves, precision-recall curves, and ablation study for its strength and efficacy.

The results of this study show that the proposed framework is a reliable and computationally efficient computer-aided diagnostic system for the automatic screening of retinal diseases. This approach can assist ophthalmologists in the clinical decision-making process, reduce the diagnostic effort, and enable large-scale screening plans, especially in resource-constrained healthcare settings, by improving the classification of both majority and minority disease classes. The positive results of the proposed framework open several directions for future research. Future studies will focus on assessing the proposed model with bigger, multicentre datasets of retinal images for further improving its generalization potential among different patient populations. Furthermore, improved imbalance learning algorithms, attention-based architectures, Vision Transformers (ViTs), and ensemble deep learning models can be explored to further improve classification performance. The use of understandable artificial intelligence (XAI) methods such as Grad-CAM, SHAP, or LIME can give visual descriptions for the model predictions, and hence, improve interpretability and increase physician confidence in automated diagnostic systems. Besides, the proposed methodology may allow its application for real clinical screening if it can be implemented in early

diabetic retinopathy detection and prevention of diabetes-induced vision loss.

FUNDING

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

COMPETING INTEREST

The authors declare that there are no conflicts of interest related to this manuscript. No financial, professional, or personal relationships with organizations or individuals have influenced the research, analysis, or conclusions presented in this work.

ACKNOWLEDGMENT

The authors would like to thank the management of Sathyabama Institute of Science and Technology for encouraging us to carry out this research.

REFERENCES

- [1] D. S. W. Ting, G. C. M. Cheung, and T. Y. Wong, "Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges," *Clin. Exp. Ophthalmol.*, vol. 44, pp. 260–277, 2016, doi: 10.1111/ceo.12696.
- [2] V. Gulshan, L. Peng, M. Coram, et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, pp. 2402–2410, 2016, doi: 10.1001/jama.2016.17216.
- [3] M. Shoruffzaman, M. S. Hossain, and A. El Saddik, "An explainable deep learning ensemble model for robust diagnosis of diabetic retinopathy grading," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 17, no. 3s, pp. 1–24, 2021.
- [4] M. Prajapati, S. K. Baliarsingh, and J. Hota, "MobileNetV3-based classification model for diabetic retinopathy," in *Proc. 10th Int. Conf. Signal Process. Integrated Netw. (SPIN)*, 2023, pp. 213–217, doi: 10.1109/SPIN57001.2023.10117341.
- [5] T. Yang et al., "Cross-modality transfer learning with knowledge infusion for diabetic retinopathy grading," *Front. Med.*, vol. 11, Art. no. 1400137, 2024, doi: 10.3389/fmed.2024.1400137.
- [6] H. Zhao, X. Li, and H. Chen, "FundusGAN: Fundus image synthesis based on semi-supervised learning for diabetic retinopathy and age-related macular degeneration," *Biomed. Signal Process. Control*, vol. 86, Art. no. 105403, 2023.
- [7] E. S. Cutur and N. G. Inan, "Multi-class classification of retinal eye diseases from ophthalmoscopy images using transfer learning-based vision transformers," *J. Imaging Informatics Med.*, vol. 38, no. 5, pp. 3110–3124, 2025, doi: 10.1007/s10278-025-01416-7.
- [8] K. Zhang, Y. Guo, and X. Wang, "Federated learning for privacy-preserving diabetic retinopathy detection," *IEEE J. Biomed. Health Inform.*, vol. 27, no. 6, pp. 2799–2807, Jun. 2023, doi: 10.1109/JBHI.2023.3252297.
- [9] K. Rana, J.-H. Ahn, and H. Kim, "Attention-based ensemble deep learning for diabetic retinopathy classification and lesion localization," *Comput. Biol. Med.*, vol. 165, Art. no. 107505, 2023, doi: 10.1016/j.combiomed.2023.107505.
- [10] P. Afshar, A. Mohammadi, K. N. Plataniotis, and A. Oikonomou, "Capsule networks for biomedical image analysis: A review and future perspectives with diabetic retinopathy case study," *Pattern Recognit. Lett.*, vol. 162, pp. 1–9, 2022, doi: 10.1016/j.patrec.2022.10.013.
- [11] Y. Wang, S. Wang, and Y. Zhang, "Dual-path convolutional neural network for simultaneous detection of exudates and hemorrhages in color fundus images," *Comput. Biol. Med.*, vol. 153, Art. no. 106427, 2023, doi: 10.1016/j.combiomed.2022.106427.
- [12] Z. Li, S. Keel, C. Liu, Y. He, W. Meng, J. Scheetz, P. Y. Lee, and D. S. W. Ting, "An automated grading system for detection of vision-threatening referable diabetic retinopathy based on color fundus photographs," *Diabetes Care*, vol. 42, no. 7, pp. 1349–1356, 2019, doi: 10.2337/dc18-0148.
- [13] A. M. Mutawa, M. Abdar, U. R. Acharya, and T. Makombe, "Application of transfer learning with deep convolutional neural networks for automated diabetic retinopathy detection and grading: A comparative study," *Appl. Sci.*, vol. 13, no. 9, Art. no. 5685, 2023, doi: 10.3390/app13095685.
- [14] M. R. Shoaib, S. M. Anwar, A. Javed, and M. A. Khan, "DiaCNN: An automated system for diabetic retinopathy detection using transfer learning and deep convolutional neural networks," *Comput. Biol. Med.*, vol. 167, Art. no. 107676, 2024, doi: 10.1016/j.combiomed.2024.107676.
- [15] V. Chaurasia, N. K. Verma, and P. Tomar, "Ensemble-based deep learning model for diabetic retinopathy detection," *Biomed. Signal Process. Control*, vol. 87, Art. no. 104923, 2023, doi: 10.1016/j.bspc.2023.104923.
- [16] K. Pravin, M. S. Nair, and T. R. Babu, "Multi-class diabetic retinopathy grading using DenseNet with image pre-processing and dense feature propagation," *Comput. Biol. Med.*, vol. 157, Art. no. 106771, 2023, doi: 10.1016/j.combiomed.2023.106771.
- [17] N. T. Abdelhamid, A. A. El-Sawy, and A. M. El-Sawy, "Deep transfer learning for diabetic retinopathy recognition using optical coherence tomography images," *Alexandria Eng. J.*, vol. 83, pp. 663–677, 2024, doi: 10.1016/j.aej.2023.08.012.
- [18] S. Ahmed, M. Paul, and K. Roy, "Multi-class classification of diabetic retinopathy using transfer learning with NASNet-Large and ResNet-50 on APTOS 2019 dataset," *Biomed. Signal Process. Control*, vol. 87, Art. no. 104903, 2024, doi: 10.1016/j.bspc.2023.104903.
- [19] M. Talukder, M. T. Islam, and M. A. Kabir, "Ensemble transfer learning approach for diabetic retinopathy detection using concatenated deep features," *Comput. Biol. Med.*, vol. 160, Art. no. 106909, 2023, doi: 10.1016/j.combiomed.2023.106909.
- [20] S. Ghosh and S. Chatterjee, "An ensemble approach for diabetic retinopathy classification using deep convolutional neural networks," *Comput. Biol. Med.*, vol. 155, Art. no. 106427, 2023.
- [21] A. Hannan, Z. Mahmood, R. Qureshi, and H. Ali, "Enhancing diabetic retinopathy classification accuracy through dual-attention mechanism in deep learning," *Comput. Methods Biomech. Biomed. Eng.: Imaging Vis.*, vol. 13, no. 1, Art. no. 2539079, 2025.
- [22] A. Alzubaidi, M. Alsharif, and A. Alwabel, "Automatic detection of diabetic hypertensive retinopathy in fundus images using transfer learning," *Appl. Sci.*, vol. 13, no. 8, Art. no. 4695, 2023.
- [23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2015.
- [24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [25] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.
- [26] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 4510–4520.
- [27] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2019, pp. 6105–6114.
- [28] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 4700–4708.
- [29] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.
- [30] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.

- [31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. Int. Conf. Learn. Representations (ICLR), 2015.
- [32] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," PLoS ONE, vol. 10, no. 3, Art. no. e0118432, 2015, doi: 10.1371/journal.pone.0118432.
- [33] X. Wei, Y. Liu, F. Zhang, L. Geng, C. Shan, X. Cao, and Z. Xiao, "MSTNet: Multi-scale spatial-aware transformer with multi-instance learning for diabetic retinopathy classification," Medical Image Analysis, vol. 102, Art. no. 103511, 2025, doi: 10.1016/j.media.2025.103511.
- [34] A. Bilal et al., "Quantum chimp-enhanced SqueezeNet for precise diabetic retinopathy classification," Scientific Reports, vol. 15, Art. no. 12890, 2025, doi: 10.1038/s41598-025-97686-w.
- [35] Y. Kaya, "A novel method for optic disc detection in retinal images using the cuckoo search algorithm and structural similarity index," Multimedia Tools and Applications, vol. 79, nos. 31–32, pp. 23387–23400, 2020, doi: 10.1007/s11042-020-09080-5