Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2026 (44) 2** : 1–7.
ISSN-0037-8712
doi:10.5269/bspm.80579

# Enhancing Diabetic Retinopathy Detection with a Hybrid Model of DenseNet121 and Support Vector Machines

Vijaya Kumar Sagenela, Karthika Gidijala, G. Neeraja Rani, Santoshachandra Rao Karanam, P. Nagamani and Naresh Tangudu

ABSTRACT: Diabetic Retinopathy is the most prevalent and common diabetes-linked ocular complication that remains one of the leading causes of vision impairment worldwide. The disease develops as a result of chronic hyperglycemia, which further promotes the development of retinal microvasculature, leading to eventual vision loss if not detected and addressed in a timely manner. In this case, it is essential to establish an effective and accurate diagnostic framework to facilitate the early detection and prevention of blindness. This paper outlines a hybrid framework for the classification of DR which integrates DenseNet121's ability to extract features and the Machine learning group's inherent classification with Support Vector Machines to diagnose DR into five phases: No DR, Mild, Moderate, Severe, and Proliferative DR. The presented model appropriately manages the most common issues associated with classification, including class imbalance, image noise, and imaging acquisition variance using machine learning pre-processing strategies and data augmentation. For the performance of the ten most popular pre-trained convolutional neural network architectures, a Stratified Whole Dataset was applied, conditioned on the source of the image: IDRiD, Kaggle, and AEH. DenseNet121 was identified as the best-performing pre-trained architecture, exhibiting a high sensitivity of 93.169 and an F1-score of 87.726, ahead of EfficientNetB0 and Xception. Taken together, the results provide strong evidence that strong pre-processing and model selection can significantly improve classification accuracy. The proposed methodology offers a valid, scalable, and patient-centered technique for self-regulating retinal image evaluation, possessing considerable potential for application in patient care. Subsequent experiments will be directed to increasing the performance of the models with suboptimal performance and the Ensemble learning technique and fulfill the technique verification over bigger and more varied datasets for practical application.

Keywords: Diabetic Retinopathy (DR), hybrid model, DenseNet121, Support Vector Machine (SVM), medical image classification, deep learning, computer-aided diagnosis, convolutional neural networks.

## Contents



## 1. Introduction

Ophthalmologically, DR is a serious microvascular consequence of diabetes mellitus, and yet remains one of the prime causes of avoidable blindness in the globe. The prevalence of DR is increasing rapidly worldwide because it is estimated that more than 463 million people have been diagnosed with diabetes. In order to avoid irreversible vision loss, early detection with the precise classification of DR stages such as No DR, Mild Non-Proliferative Diabetic Retinopathy (NPDR), Moderate NPDR, Severe NPDR, and Proliferative Diabetic Retinopathy (PDR) is critical. Ophthalmologists' traditional manual grading of retinal fundus images is time-consuming, resource-intensive, and susceptible to inter-observer variability, particularly in countries with limited resources. An automated approach to medical image analysis

2020 *Mathematics Subject Classification:* 68T07, 68T45, 92C55, 62H30, 68U10.
Submitted December 04, 2025. Published March 14, 2026

has significantly transformed the healthcare sector, driven by recent developments in deep learning and artificial intelligence (AI) [1] [2] [3] [4] [14] [15] [16]. Among these, hybrid models that combine multiple machine learning techniques have emerged as a promising approach to improve diagnostic accuracy and efficiency. The proposed hybrid model utilizes state-of-the-art deep learning architectures, such as DenseNet121, to extract features from the input fundus images and classify them with Support Vector Machines (SVM) for robust classification. Improved image quality, diversity, and reliability of medical image datasets are primarily achieved through pre-processing and augmentation of the images. These pre-processing and augmentation layers not only restore incorrectly illuminated, oriented, and contrasted images but also increase the number of images generated through artificial transformations, such as rotation, elastic deformation, and photometric transformations. After the data was prepared, we used the pre-trained DenseNet121 model to extract deep and unique features from the improved fundus images.

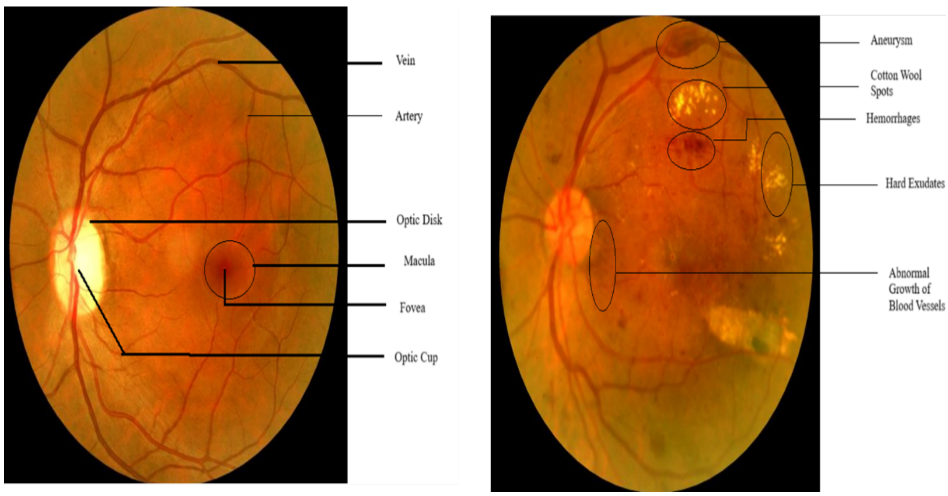


Figure 1: Normal vs. Diabetic Retinopathy Fundus Image

These features are then fed into the Support Vector Machine classifier, which is able to classify images as one of the five classes tested. No DR, Mild, Moderate, Severe motion, or Proliferative Diabetic Retinopathy. It is clearly evident that the normal fundus image and the pathological fundus image with the signs of diabetic retinopathy can be easily distinguished using the described model, as shown in Figure 1 below. Figure 1: Illustration of the differences uncovered using this model and improved fundus imaging from pathological fundus imaging due to Diabetic Retinopathy.

In summary, the proposed work produces a systematic framework of DR detection. This step involves image pre-treatment, patient training, and assessing the classification model. Shown in Figure 2, this work applied a set of images from public benchmarking data and obtained high model accuracy. The model's transparency and speed make it well-suited for the forthcoming accessible retinal screening model, provided that the correct steps are taken to address the problematic cases of low-quality images, insufficient data, and identifying lesions that require patient eye care. [9][10][11][12][19][20].

## 2. Methodology

The workflow of the proposed hybrid framework for integrated feature extraction and classification in diabetic retinopathy detection consists of the following key steps.

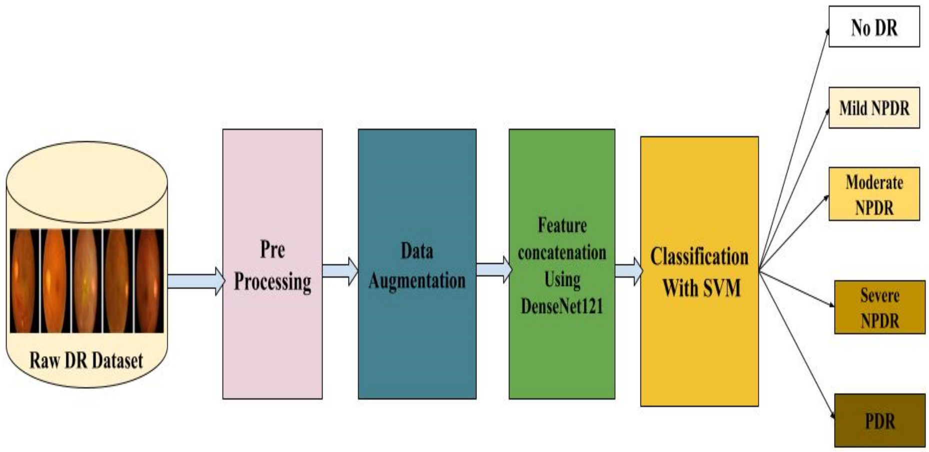


Figure 2: Workflow of the Integrated Feature Extraction and Classification of Diabetic Retinopathy

### 2.1. Data Pre-processing

Pre-processing refers to the enhancement of the raw fundus images' quality and their smoother analytical preparation for feature extraction and classification. Specifically, it is meant that grayscale conversion reduces image data to a single intensity channel, erasing all details except the main retinal ones. Secondly, normalization is a process in which pixel intensity values in essential datasets need to be scaled in a way that increases and maximizes brightness and contrast. Lastly, artifact removal represents the process of thresholding and contouring that eliminates unnecessary background elements from the images, leaving the retinal data intact.[13] [4].

### 2.2. Data Augmentation

In summary, multiple image augmentation techniques mitigate data imbalance and contribute to the model's generality. Geometric transformations further reduce spatial variability and mitigate overfitting during training, as achieved by applying random rotations of up to $\pm 20$ degrees and scaling the images. In addition, the elastic deformations, as a natural variability in the images and similar distortions that occur in fundus images when captured from different retinal centers, generate a robust model. Moreover, photometric variations, which randomly change the intensity and contrast of the image, introduce variability to the framework, allowing the model to be trained with multiple imaging settings. [15][16].

### 2.3. Feature Extraction

A DenseNet121 model, a deep CNN already pre-trained on the ImageNet dataset, is utilized to extract the CNN features from the fundus images. The network's classification layer is removed and the output of the global average pooling layer is employed to generate rich high-dimensional feature illustrations. The feature vectors encapsulate critical structural and textural intelligence of the fundus images, which form the basis for accurate classification and optimal analysis.

### 2.4. Feature Concatenation

The feature vectors obtained from several augmentations of the same image are concatenated. This way, a global representation of the data is generated so that a robust and complete model can be developed to understand the small patterns in DR classification

### 2.5. Classification Using SVM

The concatenated feature vectors are put into the SVM classifier. The particular type of SVM is a suitable choice here as the SVM classifier has high dimensionality and strength. Then the SVM model is trained to classify the fundus images to 5 possible classes: No DR, Mild NPDR, Moderate NPDR, Severe NPDR, and PDR.

**Algorithm for Hybrid Model for DR Classification**

Input: Raw DR Dataset $\mathcal{D} = \{x_i, y_i\}_{i=1}^{N}$, where $x_i$ represents the raw image, and $y_i$ represents the label (No DR, Mild, Moderate, Severe, or PDR). Output: Predicted DR class labels $\widehat{y_i}$. Steps:

a. Preprocessing: Convert the raw image $x_i$ into grayscale:

$$I_{gray}(x_i) = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$$

- Perform normalization to rescale pixel values:

$$I_{norm}(x_i) = \frac{I_{gray}(x_i) - \mu}{\sigma}$$

- Where $\mu$ and $\sigma$ are pixel intensities' mean and standard deviation. Remove artifacts using thresholding:

$$I_{binary}(x_i) = \begin{cases} 255 & \text{if } I_{gray}(x_i) > \text{Threshold} \\ 0 & \text{otherwise} \end{cases}$$

b. Data Augmentation: Apply geometric transformations:

$$T_{rotate}(x_i) = \text{Rotate}(x_i, \theta)$$

- Where $\theta$ is the rotation angle. Elastic deformation:

$$T_{elastic}(x_i) = G(x_i, \alpha, \sigma)$$

- Where $G$ applies random elastic transformations with parameters $\alpha$ and $\sigma$. Brightness and contrast adjustment:

$$T_{photometric}(x_i) = \text{Adjust}(x_i,\ \text{brightness},\ \text{contrast})$$

c. Feature Extraction: Use DenseNet121 to extract features from preprocessed and augmented images:

$$\mathcal{F}(x_i) = \text{DenseNet121}(x_i; \theta_{\text{DenseNet}})$$

- Where $\theta_{\text{DenseNet}}$ represents the pre-trained weights of the DenseNet121 model. 4. Feature Concatenation: Concatenate the extracted features to form a combined feature vector:

$$\mathcal{F}_{concat}(x_i) = \text{Concat}([\mathcal{F}_1(x_i), \mathcal{F}_2(x_i), \ldots, \mathcal{F}_k(x_i)])$$

e. Classification Using SVM: Train a Support Vector Machine (SVM) classifier with concatenated features:

$$\widehat{y_i} = \text{argmax}_c (w_c \cdot \mathcal{F}_{concat}(x_i) + b_c)$$

- Where $w_c$ and $b_c$ are the weights and bias for class $c$.

f. Output: Assign the predicted label $\widehat{y_i}$ from the SVM output.

Support vector Machine (SVM) classifier with concatenated features:

$$\widehat{y_i} = \text{argmax}_c \left( w_c \cdot \mathcal{F}_{concat} \left( x_i \right) + b_c \right)$$

Where $w_c$ and $b_c$ are the weights and bias for class $c$. 6. Output: Assign the predicted label $\widehat{y_i}$ from the SVM output.

## 3. Results and Discussion

Table 1 summarizes the comparative performance of CNN models on the dataset. Reporting the outcomes on accuracy, precision, recall, and F1-score provides the primary metrics for evaluating classification performances. DenseNet121 shows the best performance, with an accuracy of 92.83, precision of 88.73, recall of 87.04, and F1-score of 87.73. These outcomes confirm that the model can accurately extract discriminative features for the classification of diabetic retinopathy. EfficientNetB0 and Xception slightly have lower accuracy at 91.27 and 92.72, respectively. Their lower performance is due to the design of these models being relatively more depth or width or computational level. These models' width, depth, and structure have been exactly balanced to attain higher accuracy and precision. ResNet50 and MobileNetV2 have an accuracy of 33.68 and 44.18, respectively. These networks appear not well detect margins and accurately say how to make the classifier of these images, as they were not exactly able to detect margins. NASNetMobile has achieved an accuracy of 66.06; it has a low F1-score of 33.08, and it is evident that NASNetMobile struggles with an imbalance dataset and the different variations in retinal structures. In summary, the outcomes confirm that DenseNet121 is an effective multimodal model for the detection of diabetic retinopathy. The dens connected layer defines the propagation of reused and more representative features as a trivially process to tackle variations in image-level as well as training error. On the other hand, the efficient performance of EfficientNetB0 and Xception complements the process of tuning the feature miner as "... the need for tuning the feature miner is of utmost importance in Medical Imaging". The models, ResNet50, MobileNetV2, and NASNetMobile, which both awarded moderate results, guide that enhanced data balance, tunning, and hybridization techniques are demanding future work. It means choosing a modern model architecture that is perfect for the complexity dataset and for the research to work with high accuracy. The effects of proper pre-processing, augmentation, and hybrid techniques have eventually led to precise classification. The future studies will use high models to achieve greater accuracy, and we will even implement an ensemble model. It will boost distinct variance and robustness to be employed on the system.

| Model | Epochs | Batch Size | Learning Rate | Accuracy | Precision | Sensitivity (Recall) | F1-Score |
|---|---|---|---|---|---|---|---|
| EfficientNetV2S | 7 | 32 | 0.0001 | 0.879418 | 0.829202 | 0.778601 | 0.797925 |
| EfficientNetB0 | 10 | 32 | 0.0001 | 0.912682 | 0.863937 | 0.842214 | 0.851198 |
| ResNet50 | 5 | 32 | 0.0001 | 0.336798 | 0.248407 | 0.326699 | 0.224059 |
| **DenseNet121** | **10** | **32** | **0.0001** | **0.928274** | **0.887315** | **0.87037** | **0.877326** |
| MobileNetV2 | 4 | 32 | 0.0001 | 0.441788 | 0.396774 | 0.207049 | 0.138834 |
| InceptionV3 | 10 | 32 | 0.0001 | 0.887734 | 0.816351 | 0.833231 | 0.823043 |
| Xception | 7 | 32 | 0.0001 | 0.927235 | 0.892373 | 0.868047 | 0.879487 |
| VGG16 | 10 | 32 | 0.0001 | 0.804574 | 0.739423 | 0.581217 | 0.599912 |
| VGG19 | 10 | 32 | 0.0001 | 0.800416 | 0.71613 | 0.581397 | 0.592384 |
| NASNetMobile | 6 | 32 | 0.0001 | 0.606029 | 0.314861 | 0.36568 | 0.330868 |

Table 1: Performance of Pre-Trained Networks on In-House DR Dataset (IDRiD, Kaggle, AEH)

## 4. Conclusion

In general, this work introduced a hybrid model for DR classification, which combined DenseNet121 feature extraction skills and SVM classification competence. The framework thereby appeared to be beneficial for desired DR classification tasks and tackling current challenges, such as data imbalance and high noise in medical images and rather successful inference from them, acquiring from the state-of-the-art in pre-processing, data augmentation, and pre-trained networks. Particularly, examining ten pre-trained models against the internal dataset, it was revealed that DenseNet121 turned out to be a clear leader with 92.83% accuracy and 87.73% f1 -score. It should be also mentioned that EfficientNetB0 and Xception also seem to be extremely effective for this designation. Therefore, the model confirmed the crucial importance of both architectures but total data augmentation for the final classification. Moreover, even if the models like ResNet50 and MobileNetV2 work poorly, the authors received meaningful insights as for the architectural affordances and the characteristics of the dataset. The work can clearly demonstrate the potential use of hybrid solutions within medical imaging, yielding possible methods for scalable and accurate DR classification. Consequently, the research also illustrates further research towards the application of integration ensemble methods, enhancement of worse-performing models, and tests on additional datasets to make this approach general and truly applicable. Thus, the introduced research presents a significant contribution to the development of AI with the view in healthcare, making accessible and accurate DR-screening tools available worldwide.

## References

1. World Health Organization, “Diabetes prevalence to affect 783 million individuals by 2045,” *WHO*, 2024. [Online]. Available: https://www.who.int/diabetes. [Accessed: Nov. 27, 2024].

2. A. S. Brar, J. Sahoo, U. C. Behera, J. B. Jonas, S. Sivaprasad, and T. Das, “Prevalence of diabetic retinopathy in urban and rural India: A systematic review and meta-analysis,” *Indian Journal of Ophthalmology*, vol. 70, no. 6, pp. 1945–1955, Jun. 2022. DOI: 10.4103/ijo.IJO_2206_21.

3. 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.

4. V. Venugopal, J. Joseph, M. Vipin Das, and M. K. Nath, “An EfficientNet-based modified sigmoid transform for enhancing dermatological macro-images of melanoma and nevi skin lesions,” *Computer Methods and Programs in Biomedicine*, vol. 222, 106935, 2022. DOI: 10.1016/j.cmpb.2022.106935.

5. T. Brinker, A. Hekler, J. Utikal, N. Grabe, D. Schadendorf, J. Klode, C. Berking, T. Steeb, A. Enk, and V. Kalle, “Skin cancer classification using convolutional neural networks: Systematic review,” *Journal of Medical Internet Research*, vol. 20, no. 10, 2018.

6. J. Lameski, A. Jovanov, E. Zdravevski, P. Lameski, and S. Gievska, “Skin lesion segmentation with deep learning,” *Proc. Conf.*, 2019, pp. 1–5.

7. B. Harangi, “Skin lesion classification with ensembles of deep convolutional neural networks,” *Journal of Biomedical Informatics*, vol. 86, 2018.

8. L. Zheng, Z. Wang, J. Liang, S. Luo, and S. Tian, “Effective compression and classification of ECG arrhythmia by singular value decomposition,” *Biomedical Engineering Advances*, vol. 2, 100013, 2021. DOI: 10.1016/j.bea.2021.100013.

9. T. S. Roy, J. K. Roy, and N. Mandal, “Classifier identification using deep learning and machine learning algorithms for the detection of valvular heart diseases,” *Biomedical Engineering Advances*, vol. 3, 100035, 2022. DOI: 10.1016/j.bea.2022.100035.

10. C. Vasilakos, D. Kavroudakis, and A. Georganta, “Machine learning classification ensemble of multitemporal Sentinel-2 images: The case of a mixed Mediterranean ecosystem,” *Remote Sensing*, vol. 12, no. 12, p. 2005, 2020.

11. D. Tarasewicz, A. J. Karter, N. Pimentel, H. H. Moffet, K. K. Thai, D. Schlessinger, O. Sofrygin, and R. B. Melles, “Development and validation of a diabetic retinopathy risk stratification algorithm,” *Diabetes Care*, vol. 46, no. 5, pp. 1068–1075, May 2023.

12. C. L. Lin and K. C. Wu, “Development of revised ResNet-50 for diabetic retinopathy detection,” *BMC Bioinformatics*, vol. 24, no. 1, p. 157, Apr. 2023. DOI: 10.1186/s12859-023-05293-1.

13. H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, “Convolutional neural networks for diabetic retinopathy,” *Procedia Computer Science*, vol. 90, pp. 200–205, 2016. DOI: 10.1016/j.procs.2016.07.014.

14. M. D. Abramoff *et al.*, “Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning,” *Investigative Ophthalmology & Visual Science*, vol. 57, no. 13, pp. 5200–5206, 2016.

15. X. Wang, Y. Lu, Y. Wang, and W. B. Chen, “Diabetic retinopathy stage classification using convolutional neural networks,” *Proc. Int. Conf. Information Reuse and Integration for Data Science*, 2018, pp. 465–471.

16. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *Proc. Int. Conf. Learning Representations (ICLR)*, 2015.

17. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, 2017.

18. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.

19. Kaggle Dataset, “Diabetic Retinopathy Detection,” [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection.

20. A. Samanta, A. Saha, S. C. Satapathy, S. L. Fernandes, and Y. Zhang, “Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset,” *Pattern Recognition Letters*, vol. 135, pp. 293–298, 2020. DOI: 10.1016/j.patrec.2020.04.026.

21. E. Benmalek, J. Elmhamdi, and A. Jilbab, “Comparing CT scan and chest X-ray imaging for COVID-19 diagnosis,” *Biomedical Engineering Advances*, vol. 1, 100003, 2021. DOI: 10.1016/j.bea.2021.100003.

*Karthika Gidijala (Corresponding author),*
*Assistant Professor, Department of Computer Science and Engineering,*
*GITAM (Deemed to be University), Hyderabad, India.*
*E-mail address:* kgidijal@gitam.edu

*and*

*Vijaya Kumar Sagenela,*
*Assistant Professor, Department of Computer Science and Engineering,*
*GITAM (Deemed to be University), Hyderabad, India.*

*and*

*G. Neeraja Rani,*
*Geethanjali College of Engineering and Technology,*
*Cheeryal, Hyderabad, India.*

*and*

*Santoshachandra Rao Karanam,*
*Assistant Professor, Department of Computer Science and Engineering,*
*GITAM (Deemed to be University), Hyderabad, India.*

*and*

*P. Nagamani,*
*Department of Information Technology,*
*Anurag University, Hyderabad, India.*

*and*

*Naresh Tangudu,*
*Department of Computer Science and Engineering (Data Science),*
*Aditya Institute of Technology and Management, Tekkali, India.*