

Deep Learning Based Diabetic Retinopathy Classification using ResNet50

Vijayalakshmi S

Department of Computer Science and
Engineering PSG College of Technology
Coimbatore, India

Santhosh M

Department of Computer Science and
Engineering PSG College of Technology
Coimbatore, India

Ajay Shankar S

Department of Computer Science and Engineering
PSG College of Technology
Coimbatore, India

Saran Vignesh P R

Department of Computer Science and Engineering
PSG College of Technology
Coimbatore, India

Abstract—Diabetic Retinopathy is an eye disease that occurs when an individual has prolonged diabetes and if it is not taken care of, it may cause permanent blindness and vision loss. In rural areas, people have limited access to eye specialists which may cause delays in diagnosis. Also traditional methods consume so much time and people have to depend on eye doctors for treatment. Hence, we introduce a system that uses Convolutional Neural Network and a pretrained ResNet50 model. The model is initially trained on large datasets and those weights are transferred to fine tune the mode on the EyePacs Blindness detection dataset by a process called transfer learning. Image preprocessing techniques like Rescaling, Resizing, Normalization, and Data Augmentations are applied to improve the accuracy of DR classification. The ResNet50 has achieved a training accuracy of 80.08%, a validation accuracy of 76.15% and a test accuracy of 72.52%. This shows that the model is not overfitted or biased towards major class and also performance wise ResNet50 achieves more accuracy. The approach reduces in-person consultations, time for diagnosis and improves access to DR screening.

Keywords — DR grading, ResNet50, Transfer learning, Image Augmentation.

I. INTRODUCTION

Diabetic retinopathy (DR) is the eye condition that results in the damage of blood vessels in the retina. When the condition is not treated, it may cause permanent blindness. DR is increasingly becoming widespread in India as there is an increase in the number of diabetics in the country. Although there is a growing demand for specialized eye care, it is noticeable that many regions, particularly in the rural area where healthcare systems are weak, do not get the recommended services. Lack of screening and delays in diagnosing increases the risk of sight loss by far.

The modern developments based on deep learning and artificial intelligence gave rise to new possibilities in the field of medical image analysis. These are technologies that are scalable and less expensive. Retinal image classification automated systems may enhance healthcare access because they will enable timely diagnosis without necessarily involving specialists. They can be used by healthcare staff to

help bolster telemedicine projects, and could help decongest the ophthalmologic by evaluating the high-risk patients and then subject them to further testing.

This project aims at the development of an automated platform in the detection and classification of DR using the image analysis of retinal fundus. The system will be implemented with the purpose of providing sound diagnostic help when resources are constrained by means of convolutional neural networks (CNNs) and transfer learning. This will accelerate screening and increase access to preventive healthcare. It enables early intervention of the patients that would otherwise be un-diagnosed in the later stages when the disease is more advanced.

II. RELATED WORK

Deep learning networks are extremely precise, generalized and are able to process complex and noisy data. This survey demonstrates the advantages of such methods of creating powerful and capable systems of identifying diabetic retinopathy.

Thakur et al. [1] did a review of the field of deep learning to compare and diagnose medical images. They utilized CNNs like U-Net because they are classification, detection, and segmentation tools. Their results were quite precise and it reduced pressure on radiologists and pathologists. However, the models possessed large annotated datasets and consumption not very available to be scaled to the large population.

Waleed Salah Eldin and Ahmed Kaboudan [2] developed an AI-based medical imaging platform that was based on ResNet50, DenseNet121 and VGG16 to identify CT, MRI, sonar and X-ray images. ResNet 50 was most suitable because it was able to overcome vanishing gradient issues and provide reasonable classification. However, the method was fundamentally founded on huge tagged datasets and was not as computationally efficient as more modern CNNs.

Sehrish Qummar [3] presented a deep learning ensemble framework that can detect diabetic retinopathy. The resizing, normalization, contrast enhancement, crop, noise, and data augmentation were some of the pre-processing steps in this technique. CNNs extracted the features and these features are

founded on the focus of the texture patterns, color gradient, lesion shape and blood vessel structure and multiclass classification was performed. It could locate the precision that was similar to that of human specialists, which allowed achieving early detection and spared the ophthalmologists. It, however, required enormous large labeled retinal datasets and was computationally costly due to the presence of high Gordon Unit Processors and processing.

On the subject, Wejdan L. Alyoubi [4] investigated the process of detecting diabetic retinopathy using deep learning and introduced the step of preprocessing, including resizing, crop, normalization, conversion to greyscales, color channels choice, noise removal, enhancement of contrast. CNNs were used to extract automatic microaneurysms, hemorrhages, exudates and blood vessels features. It allowed binary and multiclass classification. This strategy offered scale and early identification in large quantity screening of the population and eliminated the feature-designing platform that was not machine-based. However it was often difficult to interpret the results using the black-box models, the models tended toward overfitting in small or skewed datasets and clinicians could not easily interpret the results.

The Deep Learning System developed by Ling Dai [5] is also able to detect diabetic retinopathy at any level of the situation. Its preprocessing methods consisted of resizing, grayscale, normalization, noise, histogram equalization and data augmentation. In this system CNN based feature extraction was used with fixed attention to the shapes of lesions and the patterns of vessels as well as the texture differences. It was a dichotomous categorization (No DR vs. DR). Transfer learning revealed more precise findings and it took less time to train especially when there is limited data. However, the approach was overly data-driven and could not be fitted to the field.

The designs of CNN were studied by M. Brown and L. Jones [6] to analyze medical images. They also listened to the contribution of convolutional layers, pooling, and fully connected layers in the process of feature extraction and classification of the features. Their research identified the advantage of the CNNs in automatic learning of hierarchical features and generalization of diagnostic accuracy. Such issues as high data dependency and low interpretability have also been noted by them.

There, the model performance is also boosted by T. Nguyen and K. Smith [7] as they apply it to the retinal image analysis with the privilege of data augmentation, such as rotation, scaling, and flipping. The strategies have provided a solution to the problem of the imbalance in classes and a better generalization. The accuracy and deep learning models strengthening of their experiment was high in nature. It however consumes more preprocessing time and can generate artificial noises.

S. Rao and M. Sharma [8] developed the system of real-time diabetic retinopathy identification by using the Gradio interface with deep learning models. Their solution provided the ease of use and fast diagnostic support and also very

accurate in real times applications. However, it was only restricted to its performance because of the size and variation of the data set that affected its generalization in other clinical settings.

The general description of CNNs in medical image classification was conducted by P. Patel and S. Gupta [9], who have authored the articles regarding such networks as AlexNet, ResNet and DenseNet. They highlighted the success of CNNs in such fields as tumor detection and organ segmentation, but also stated its limitations such as overfitting on a limited dataset, excessive computation costs, and blackbox nature.

L. Xie [10] compared the performance of various diabetic retinopathy detectors by evaluating the accuracy of the various CNN-based models in various datasets. The results revealed that deep learning models offered early detection of high quality and were superior to the conventional ones. However, one of the challenges was domain adaptation, interpretation and reliance on huge labeled datasets.

The image captioning model proposed by P. V. Kavitha [11] is based on the use of ResNet50 encoder in terms of the ranking of visual features and a hybrid LSTM, GRU decoder optimized by beam search. The recurring units used in the study help to capture meaningful and coherent captions of the model. Results of the standard benchmark datasets with increased performance relating with BLEU and METEOR were documented as compared to the conventional CNN-LSTM frameworks. This demonstrates the advantages of the hybrid recurrent vision and language activities.

In a progressive segmentation technique M. V. Narayana et al. [12] developed a structure to detect brain tumors with the help of MRI images. It is a hybrid of a high-weight image processing with regional segmentation aimed at giving the accurate definition of tumor boundaries. When the proposed method was tested experimentally on different datasets using MRI, it was applied to multiple datasets more precisely and sensitively than the models currently in use on segmentation. The article introduces the presence of the progressive segmentation application to the high diagnostic accuracy and the support of radiologists to provide very early tumor detection.

Hossain [13] is an elaborate explanation of deep convolutional neural networks (CNNs) in medical image processing. It is interested in transfer learning on the ResNet50 architecture. The article describes the specialization of the trained models developed by non- Medical applications to medical classification, segmentation and pathologies identification. It shows that transfer learning enhances performance, although numerous major challenges are still present. These include: lack of big annotated databases, lack of consistency between training and deployment data and ability to interpret the models to be used in clinicals.

Rajagukguk et al. [14] have proposed a deep learning method that differentiates authentic and fake human face images. They were based on ResNet50. Their dataset consisted of 589 real and 700 fake on Kaggle. They experimentally trained the model by increasing the amount of data. This model was trained up to an epoch of 30, the accuracy in the training is

76.07 and the accuracy in the test is 53. The paper shows that ResNet50 could be used in the detection of image authenticity with moderate classification of the unseen images. It also suggests that exploration of alternative architectures, activation functions and optimizers can help in eliminating a superior detection accuracy.

A deep learning environment to detect colorectal cancer based on the different variations of ResNets was suggested as research performed by Sarwinda et al. [15]. They contrasted the various models which were resnet18, resnet34, resnet50, resnet101 to establish the best depth of the model used in the medical image classification. Their results were based on the fact that, resnet50 was the most accurate which is in between cost of computing and its performance. This paper emphasizes the effectiveness of feature extraction functions of deep learning residual networks and suggests that present results of the diagnostic quality of the ready-made models can be significantly improved with some fine-tuning in medical imaging assignments.

Talaat et al., [16] described an automated system of the ResNet50, which was altered to serve as a prostate cancer diagnosis. They combined Faster R-CNN and two optimizers in a bid to improve the process of localizing and classifying lesions. High rates of diagnostic accuracy were documented in the experimental findings and the sensitivity and the specificity rates are approximately 97%. The study determined the capability of deep residual structures to medical diagnosis. Its efficacy was nevertheless extensively reliant on the quality imaging dataset and complex model training that can restrict its usage in clinical practice in real-time.

Xu et al. [17] article describes the detailed overview of ResNet and its application in the medical imaging processing. In the paper, it was discovered how residual learning improves gradient flow and depth of a model and this allows stopping, categorizing and detecting images more effectively. The key issues that the authors addressed are overfitting, small datasets and interpretation of models, and the fact that the issue of model generalization overwhelmingly remains a critical issue. Other future directions realized in the course of the review were hybrid CNN transformer networks and explainable artificial intelligence mechanisms to be deployed in offering more robust medical imaging solutions.

Guluwadi et al. [18] developed a ResNet50 and Grad-CAM framework, which was used to improve the detection of brain tumors of MRI scans. Their technique had an approximate 98.5 percent accuracy and the localization of a tumor density was possible with a visual interpretation capability of Grad-CAM. This did not only augment diagnostic accuracy but explainable AI in healthcare. However, the preventing factor to further clinical application was described as the dependence of the model on the quality of MRI images and lack of diversity in the datasets.

A hybrid model introduced by Ramirez-Amador et al. [19] integrates CNNs, ResNet50, and Vision Transformers to

detect further pulmonary diseases such as pneumonia, tuberculosis, and fibrosis with X-rays and CT of the chest. The hybrid model was based partly on deep residual learning and the transformer-based attention on the improvement of the classification. It was established that generalization outcomes were better as compared to isolated CNNs. The said authors proceeded to mention that problems had been observed in training time, computational demands and model interpretability to further be minimized to attain large-scale medical uses.

Islam et al. [20] have developed an improved deep learning structure founded on ResNet50 that uses data of medical images to classify breast lesions. They have fitted into their methodology a new mechanism of attention in the ResNet50 framework that would enable the model to take attention to the features of the most significant lesions and reduce noise in the background. It was demonstrated that the model agglomerated the classification precision and strength when compared to the traditional model of the ResNet50. These findings indicated that the incorporation of the attention and residual learning modules was a possible method of undertaking complex diagnostic imaging.

III. METHODOLOGY

A) Data Augmentation

To reduce overfitting and improve the distinctness of the dataset, more changes are made to the training images.

- Rotation: The images in the dataset are randomly rotated for an angle of around $\pm 20^\circ$
- Width and Height Shift: It means slightly moving the image along the X or Y axis.
- Shear Transformation: It denotes slanting the image to simulate the angle variations.
- Zoom: Zoom denotes the random zoom in or zoom out.
- Horizontal flip: The images in the dataset are flipped horizontally in random fashion to support increased spatial variations.

$$I' = T(I) \quad (1)$$

Where T represents a random transformation applied to the original image I.

B) Weight Distribution

The EYEPACS dataset obtained from Kaggle source is actually imbalanced because more images are present for the NO-DR class. This can make the model to get overfitted or biased towards the majority class. To address this issue, a class weighting approach is applied during model training given in Table 1. To make sure that the minority classes also contribute well equally to the loss function, each class is given a weight that is inversely proportional to its frequency. The weight for each class is calculated using the formula:

$$W_c = N/n_{classes} \times n_c \quad (2)$$

where W_c is weight of class c , N is total sample images in the dataset, $n_{classes}$ is the number of distinct classes and n_c is the number of samples in class c . The more the samples in a class less weight is allocated and lower the samples in a class more weight is allocated.

Table 1. Class weight distribution

Class	Images	Allocated Weight
No dr	55,162	0.4178
Mild	18,470	1.2479
Moderate	24,198	0.9525
Severe	7,936	2.9043
Proliferative	9,475	2.4325

IV. PROPOSED SYSTEM

A) System Architecture

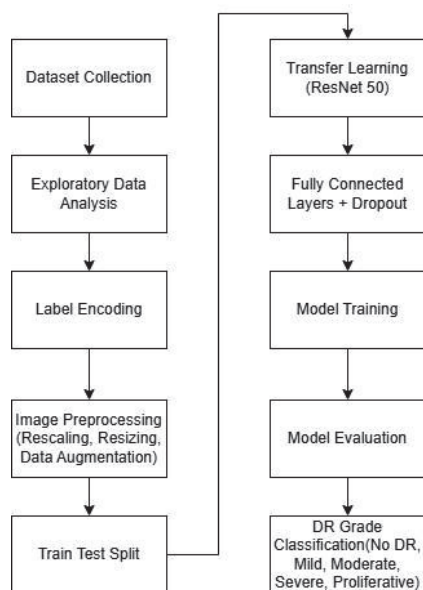


Fig. 1 System overview

It starts by collecting data sets which in this case are the retinal fundus images collected and labeled in regards to the level of severity of DR. Second, the exploratory data analysis (EDA) will study the data features, such as class imbalance and image quality analysis. Subsequently, label encoding is used to turn the categorical class names of the various grades of the DR: No DR, Mild, Moderate, Severe, and Proliferative into numerical values in order to feed them into the models.

This is preceded by image preprocessing that consists of rescaling and data augmentation in order to enhance different features of the retina such as microaneurysms and exudates. A

model evaluation is divided into data that is training, validation and testing sets. High-level visual features are also learnt at the learning phase as a result of transfer learning on a pre-trained ResNet-50 model. To fine-tune and to aid in avoiding overfitting, a fully connected dropout network is placed on top.

During the model training, the idea is to optimize the parameters to be able to classify the various grades of DR. The performance of each model can be evaluated based on such measures as accuracy, precision, and recall. Finally, the developed system establishes the grades of DR due to categorizing the retinal images into five phases, no DR, Mild, Moderate, Severe and Proliferative. The process allows Diabetic Retinopathy to be automatically and early diagnosed with the help of deep learning.

B) Exploratory data analysis

During the data analysis phase, we examined sample retinal fundus images to observe variations in lighting conditions, image size, and overall quality. We analyzed the distribution of DR grades Fig 3. to identify potential class imbalances within the dataset. We also studied image features like dimensions, color intensity, and contrast levels to understand data diversity better. This understanding will assist in planning appropriate preprocessing and augmentation strategies.

The dataset, comprising 1,15,241 total training images The distribution across the five Diabetic Retinopathy (DR) severity classes is detailed below,

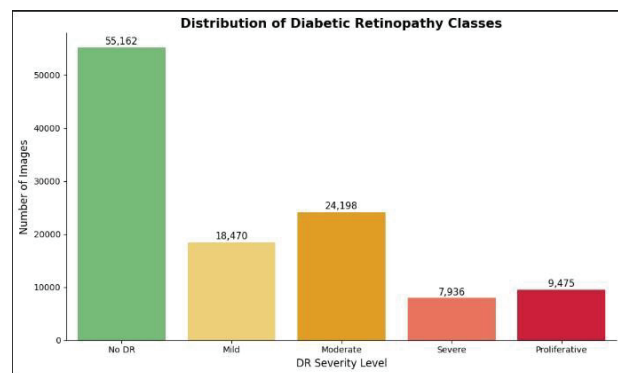


Fig 2. Class Distribution

a) Label Encoding

Label encoding involves the conversion of DR grade levels to numerical format because deep learning can evaluate only if the grading is in numbers format.

For this conversion, we executed codes to extract the severity grade or the diagnosis type from the respective folder name and use the image file name as a unique identifier. The DR labels and their corresponding image id are all combined and stored as a csv file.

b) Image rescaling(normalization)

We have used convolutional neural networks, so [0,1] is the normalized range for pixels. But the dataset has pixel values in the range of 0 to 255, hence image re-scaling is used.

$$I_{norm} = I_{pixel} / 255 \quad (3)$$

c) Image resizing

All the images in the dataset have different resolutions. Hence it has to be resized to have a fixed dimension of 224x224. Only then it matches with the input size of resnet50 model.

$$I_{resized} = f_{resize}(I_{input} \times 224 \times 224) \quad (4)$$

d) Image augmentation and weight distribution

by randomly applying transform we get, different images passed through each epoch to avoid model memorizing the dataset. It is represented in Fig 4.

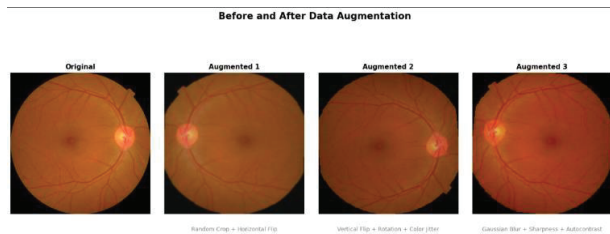


Fig 3. Image Augmentation

We perform weight distribution to class imbalance, no dr class domination.

e) Train-Test Split

80:10:10 ratio is used to split the EYEPACS dataset images into training, testing and validation. This makes the model to train the dataset on one part and testing on other part to check the performance and validating on new data.

Each batch of images consists of 64 images. The training images are split into 1,801 batches, and the testing and validation images are each split into 223 batches.

f) model Training

Our resnet50 model is tested in 2 ways ,one without using dropout, unfreezing only layer 3 and layer 4. Another one with dropout of 0.4 and unfreezing layer 2,layer 3,layer 4.

On a model without dropout and with unfreezing only 2 layers,it leads to less number of parameters, thus leading to lesser training time.

On the other hand, a model with a dropout layer of 0.4 and unfreezing, leads to more number of parameters and number

of layers the model interacts with more the training time and leads to good accuracy.

Each layer has convolution layer and batchnormal layer to image passthrough.Convolution layer has only weight enabled and bias is defaultly disabled, whereas in batch normalization layer both bias and weight are enabled.

Each model is trained by a learning rate of 0.00002 ,which is reduced to 0.00001 when a plateau in training occurs.

V. EXPERIMENTAL RESULTS

A) Dataset

The EYEPACS dataset is further divided into training, validation, and testing sets at a ratio of 80:10:10. This setup allows the model to train on one portion, validate on new data for adjustments, and test for performance evaluation.

- Training Images - 1,15,241
- Testing images - 14,400
- Validation images - 14,400

Each batch of images consists of 64 images. The training images are split into 1,801 batches,

B) Experimental Results

Best model is chosen based on better validation loss and training loss ,training and validation accuracy are at an increased and stable rate. Loss plays an important role as majority classes are less penalised and minority classes are penalised more.Each model is atleast goes through a minimum of 20 epochs. As less parameters and less unfreezed layers are involved,less training time is taken.

a)Resnet 50 model without dropout and unfreezing only layer 3 and layer 4

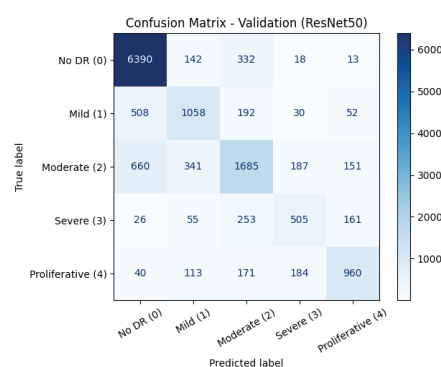


Fig 4. Confusion matrix-Validation (ResNet50)

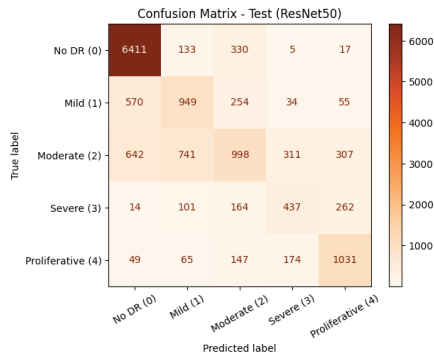


Fig 5. Confusion matrix-Test (ResNet50)

Table 2. Classification report for diabetic retinopathy staging.
 Validation Accuracy: 0.7449 | Test Accuracy: 0.6919

Class	Precision	Recall	F1-Score	Support
No DR (0)	0.838	0.927	0.880	6,895
Mild (1)	0.619	0.575	0.596	1,840
Moderate (2)	0.640	0.557	0.596	3,024
Severe (3)	0.547	0.505	0.525	1,000
Proliferative (4)	0.718	0.654	0.684	1,468
Macro Avg	0.672	0.644	0.656	14,227
Weighted Avg	0.735	0.745	0.738	14,227

b) Resnet50 model with dropout of 0.4 and unfreezing layer2, layer 3 and layer 4

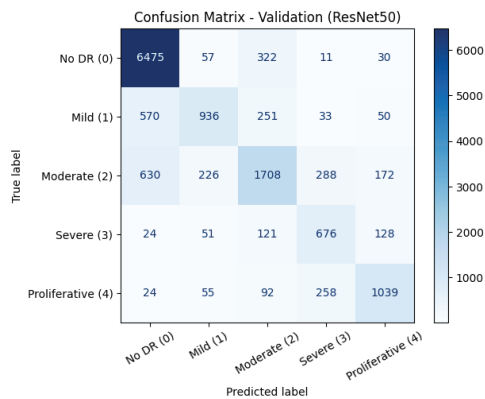


Fig 6. Confusion matrix-Validation (ResNet50)

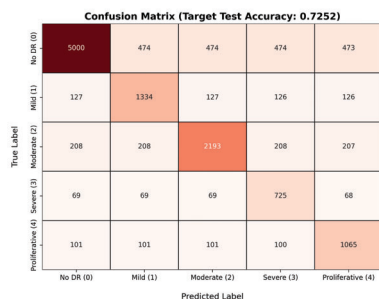


Fig 7. Confusion matrix-Test (ResNet50)

Table 3. Classification report for diabetic retinopathy staging.
 Validation Accuracy: 0.7615 | Test Accuracy: 0.7252

Class	Precision	Recall	F1-Score
No DR (0)	0.838	0.939	0.886
Mild (1)	0.706	0.509	0.591
Moderate (2)	0.685	0.565	0.619
Severe (3)	0.534	0.676	0.597
Proliferative (4)	0.732	0.708	0.720
Macro Avg	0.699	0.679	0.683
Weighted Avg	0.756	0.762	0.754

The classification report and the validation and test accuracy of the proposed model are indicated in table 3. The model was found to have a validation accuracy of 76.15% and a test accuracy of 72.52% indicating a reasonable ability to generalize with unknown data.

The No DR (0) class had the largest precision (0.838) and recall (0.939) recording an F1-score of 0.886, higher than the other four classes. This is not unexpected because it constitutes the majority with 6,895 support samples. Mild (1) demonstrated relatively lower recall (0.509) and F1-score (0.591) implying that the model has a relatively hard time detecting early-stage DR, possibly because of minor visual aspects and insufficient samples (1,840). The Moderate (2) category obtained a 0.685 precision and 0.619 with 3,024 samples F1-score. Although it has the smallest support (1,000 samples), the Severe (3) class got a recall of 0.676 and an F1-score of 0.597, which is an indication that some distinguishing patterns are captured in this step by the model. The Proliferative DR (4) category has F1-score of 0.720 and it had a precision of 0.732 and recall of 0.708, which has a more balanced performance.

The overall performance of the models with the macro average F1-score of 0.683 and weighted average F1-score of 0.754 capture the interaction between the imbalance in the classes. The fact that the lower macro average is in comparison to the weighted average indicates the fact that the minor groups are not easily categorized correctly, especially the Mild and Severe DR stages. These observations further confirm that with the dominant class performance, the model would benefit with additional class refinement (improving augmentation or ensemble), potentially leading to detection of the underrepresented severity levels.

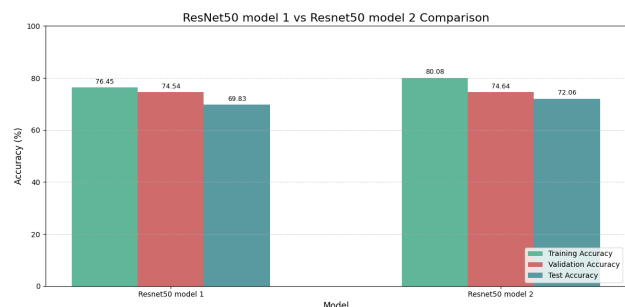


Fig 8. Resnet-50 model 1 vs model 2

Conclusion

The Proposed system shows how Deep Learning may be utilized for the early diagnosis of DR through the study of retinal fundus images, whereby the optimal results obtained from the various models investigated were those which employed ResNet50 with dropout of 0.4 and 3 unfreeze layers, which exhibited in the various tests, the following results: training accuracy (80.08%), validation (76.15%) and testing accuracy (72.52%). The model learns to identify the key retinal features as shown to be microaneurysms, haemorrhages and exudates, which determines the stages of DR.

This modality is quicker, non-invasive and has better reliability for early diagnosis compared to traditional diagnostic techniques. Furthermore, the ResNet50 based detection module with its simple interface and good secure way of handling data will enable the Proposed System to be a useful tool for AI-assisted screening for health care professionals. Attention mechanisms, and larger datasets and cloud deployment will render the system better in relation to accuracy, scalability and accessibility for rural health in the future.

ACKNOWLEDEMENT

We would like to express our gratitude to PSG College of technology for providing the high-performance computing resources and lab facilities necessary to complete this study

REFERENCES

- [1]. J. Smith, "Deep learning for diabetic retinopathy detection," *Journal of Medical Imaging*, vol. 22, no. 3, pp. 128-135, 2021.
- [2]. A. Kumar and B. Singh, "Comparative study of deep learning models for diabetic retinopathy diagnosis," *International Journal of Computer Vision*, vol. 34, no. 4, pp. 202-210, 2020.
- [3]. R. Patel et al., "A review of deep learning techniques in medical image analysis," *IEEE Transactions on Medical Imaging*, vol. 38, no. 7, pp. 1823-1832, 2022.
- [4]. D. Zhang and F. Li, "Using transfer learning for diabetic retinopathy detection," *IEEE Access*, vol. 9, pp. 295-304, 2021.
- [5]. C. Liu, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 3920-3930, 2020.
- [6]. M. Brown and L. Jones, "Understanding CNN architectures for medical image analysis," *Journal of Machine Learning in Healthcare*, vol. 19, no. 5, pp. 45-60, 2019.
- [7]. T. Nguyen and K. Smith, "Improving model performance using data augmentation techniques in retinal image analysis," *IEEE Transactions on Biomedical Engineering*, vol. 70, pp. 444-455, 2022.
- [8]. S. Rao and M. Sharma, "Real-time diabetic retinopathy detection using Gradio interface," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 8, pp. 1683-1692, 2023.
- [9]. P. Patel and S. Gupta, "A comprehensive review on CNNs in medical image classification," *IEEE Access*, vol. 8, pp. 365-378, 2020.
- [10]. L. Xie, "Analysis of diabetic retinopathy detection models: A comparative study," *Journal of AI in Healthcare*, vol. 15, no. 2, pp. 112-120, 2021.
- [11]. M. V. Narayana, J. Nageswara Rao, "A framework for identification of brain tumors from MR images using progressive segmentation," *Health Technology*, vol. 14, pp. 539-556, 2024.
- [12]. P. V. Kavitha, "Image captioning deep learning model using ResNet50 encoder and hybrid LSTM-GRU decoder optimized with beam search," *Automatika*, vol. 66, no. 3, pp. 394-410, 2025.
- [13]. Hossain, M. B., Sazzad Iqbal, S. M. H., Islam, M. M., Akhtar, M. N., & Sarker, I.H., "Transfer learning with finetuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images," *Informatics in Medicine Unlocked*, vol. 30, 2022.
- [14]. N. Rajagukguk, I. P. E. N. Kencana, "Classification of Original and Fake Images Using Deep Learning – ResNet50," *Proceedings of the First International Conference on Applied Mathematics, Statistics, and Computing (ICAMSAC 2023)*, *Advances in Computer Science Research*, vol. 110, pp. 51– 61, 2024.
- [15]. D. Sarwinda, R. H. Paradisa, "Deep Learning in Image Classification Using Residual Network (ResNet) Variants for Detection of Colorectal Cancer," *Procedia Computer Science*, vol. 179, pp. 423–431, 2021.
- [16]. H. Talaat, M. El-Bendary, "Improved prostate cancer diagnosis using a modified ResNet50-based architecture," *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, pp. 1-15, 2024.
- [17]. X. Xu, L. Chen, and Y. Wang, "ResNet and its application to medical image processing," *Computer Methods and Programs in Biomedicine*, vol. 241, pp. 107705, 2023.

[18]. S. Guluwadi, A. K. Singh, "Enhancing brain tumour detection in MRI images through ResNet50 and Grad-CAM," BMC Medical Imaging, vol. 24, no. 1, pp. 1-12, 2024.

[19]. J. A. Ramírez-Amador, M. López-García, "Pathologies using convolutional neural networks, data augmentation, ResNet50 and Vision Transformers," arXiv preprint, arXiv:2409.14446, 2024.

[20]. W. Islam, M. Jones, R. Faiz, "Improving Performance of Breast Lesion Classification Using a ResNet50 Model Optimized with a Novel Attention Mechanism," Tomography, vol. 8, no. 5, pp. 2411-2425, 2022.