INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS

Received: 08-07-2024

Accepted: 20-09-2024

Published: 20-11-2024

Citation: Velagala SR, Srinivas UM (2024) Evaluating Performance of Pre-trained Models for Diabetic Retinopathy Detection with a Minimal Dataset using Transfer Learning. Indian Journal of Science and Technology 17(42): 4395-4404. https://doi.org/10.17485/IJST/v17i42.2210

* **Corresponding author**.

sumanthvelagala@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Evaluating Performance of Pre-trained Models for Diabetic Retinopathy Detection with a Minimal Dataset using Transfer Learning

Sumanth Reddy Velagala[1]*, U Mohan Srinivas[2]

1 Student, Department of CSEAIML, Malla Reddy Engineering College, Maisammaguda, Secunderabad, Telangana, India
2 Professor, Department of CSE-AIML, Malla Reddy Engineering College, Maisammaguda, Secunderabad, Telangana, India

Abstract

Objectives: Evaluate the effectiveness of transfer learning with pre-trained CNNs for detecting diabetic retinopathy using a minimal dataset. **Methods**: Six pre-trained models (DenseNet201, ResNet152, VGG16, InceptionV3, MobileNet, and EfficientNetB0) were trained on 40 retinal fundus images each for training, validation, and testing. Models were evaluated over 5, 10, and 15 epochs, focusing on Test accuracy, Validation accuracy, and Validation loss. **Findings**: MobileNet achieved a Test accuracy of 70% at 15 epochs, consistently outperforming other models across all epochs. InceptionV3 initially performed well but was ultimately surpassed by MobileNet, which demonstrated stable and robust performance with limited data. These results suggest that MobileNet is highly effective in learning and generalizing features relevant to diabetic retinopathy detection, making it an optimal choice for use in resource-constrained settings. **Novelty**: This study highlights MobileNet's superior efficacy in diabetic retinopathy detection using minimal data, addressing an area that has been largely unexplored in existing literature.

Keywords: Diabetic retinopathy; Convolutional neural network; Transfer learning; Pretrained models; Minimal Dataset

1 Introduction

Diabetic Retinopathy (DR) is a disorder caused by the prolonged effects of diabetic mellitus[(1)], which encompasses a range of metabolic disorders characterized by sustained elevated blood sugar levels due to deficiencies in insulin secretion, action, or both, disrupting carbohydrate, lipid, and protein metabolism and frequently involving insufficient response to low insulin levels or tissue resistance to insulin[(2)]. It is indeed the most common cause of vision loss in working-age groups in the Western world[(1)]. The global prevalence of DR was (22.27%), Africa (33.8%), and Ethiopia (19.48%)[(3)]. Detection of Diabetic Retinopathy in the early stages is essential to save an individual's eyesight. The earliest symptoms of Diabetic Retinopathy are the bulging of blood vessels called Micro Aneurysms, which causes blurred vision, fluctuations in vision, and dark spots. The screening can involve examining retinal images taken with a Fundus camera, either in 2D or through OCT 3D imaging [(4)]. There has been a decline in the cost of detecting diabetic retinopathy by 1/5th of the original cost by incorporating artificial intelligence and deep learning algorithms [(5)].

Transfer learning in deep learning is a method used to train convolutional neural networks (CNN) without having to start from scratch with the neuron weights. Instead, the weights are transferred from pre-trained CNN models that were trained on large datasets, achieving high accuracy across various CNN models like ImageNet. These pre-trained weights make it easier to classify diverse datasets without needing random weight initialization. [(6)] Generally, traditional neural networks fail to classify unseen data in clinical settings, and they often fail when subjected to limited data. Therefore, transfer learning techniques offer options to utilize existing networks instead of training new ones from scratch. [(7)]

Despite the advancements in diabetic retinopathy (DR) detection, comparative analysis of pre-trained convolutional neural network models is scarce [(6)]. Particularly, the efficacy of transfer learning in a setting with a limited dataset is unexplored. Traditional neural networks require large amounts of labeled data, which are difficult to obtain and costly. This study explores the performance of six pre-trained CNN models—DenseNet201, ResNet152, VGG16, InceptionV3, MobileNet, and EfficientNetB0—in detecting DR with a minimal dataset, adjusting the number of epochs from 5 to 15. The goal is to identify the best model among the selected transfer learning options for detecting diabetic retinopathy in resource-constrained settings.

2 Literature Review

Tymchenko [(8)] developed a deep learning-based approach for detecting and classifying the stages of diabetic retinopathy (DR), aiming to improve diagnostic efficiency and consistency by leveraging convolutional neural networks (CNNs). This approach addresses the challenges posed by traditional methods, which often rely on subjective expert diagnoses. Building on previous efforts that used traditional computer vision techniques like support vector machines, deep learning methods have demonstrated superior performance in DR detection, with studies such as Pratt [(8)]. achieving high sensitivity in identifying retinal features. Further advancements in transfer learning, as seen in works by Hagos and Sarki [(8)], have enhanced the performance of CNN-based models by using pre-trained architectures like InceptionNet and ResNet50. Tymchenko [(8)]. advanced this further by proposing a multi-task learning framework, integrating classification, regression, and ordinal regression to predict DR severity. Utilizing datasets like the Kaggle APTOS 2019 and EyePACs, they fine-tuned their CNN models to achieve high sensitivity and specificity, ranking 54th out of 2,943 entries in the Kaggle competition with a kappa score of 0.925466. Their approach also employed image augmentation and an ensemble of EfficientNet and SE-ResNeXt architectures, improving model robustness. Importantly, they incorporated SHAP (Shapley Additive Explanations) to enhance model interpretability, which is vital for real-world clinical deployment. The study concludes by highlighting the potential for further improvements through hyperparameter tuning and the inclusion of other eye-related conditions, emphasizing the role of large-scale data pretraining and domain-specific fine-tuning in medical image analysis.

The review by Kim [(9)] provides an extensive analysis of transfer learning (TL) in medical image classification, emphasizing its effectiveness in overcoming challenges like data scarcity, hardware costs, and expert annotation requirements. TL, particularly through convolutional neural networks (CNNs), utilizes pre-trained models from natural image datasets (such as ImageNet) to enhance medical image classification tasks, improving performance and reducing computational demands. Reviewing 121 studies, the authors found that deep models like Inception and ResNet were commonly used due to their strong generalization across tasks, with Inception noted for its parameter efficiency, while older models like AlexNet and VGG remained in use. Feature extraction emerged as the most popular TL approach, offering computational savings, followed by fine-tuning for optimizing performance. Despite a wide variety of data modalities and medical conditions being studied, such as skin lesions and breast cancer, the review noted that larger datasets do not always correlate with improved performance. The paper also discussed recent TL advancements like unsupervised learning, GANs, and domain adaptation, while addressing limitations in fine-tuning practices. Kim [(9)]. advised starting with feature extraction and fine-tuning only when necessary, and recommended 3D CNNs for future research. The review concluded with practical guidance for selecting backbone models, advocating for deep CNNs like ResNet and Inception, and promoting the use of publicly available datasets to enhance reproducibility in medical image classification.

Bansal [(10)]. tackle the pressing issue of data scarcity in deep learning by reviewing various data augmentation techniques that address this limitation. The paper underscores the reliance of deep learning models on large datasets to avoid overfitting and enhance generalization, particularly in fields like healthcare and security, where data collection is constrained by privacy and resource challenges. The authors categorize augmentation methods into traditional approaches like geometric transformations and advanced techniques such as Generative Adversarial Networks (GANs) and neural style transfer. Highlighting strategies such as color and geometric transformations, random erasing, and GANs, the review demonstrates their efficacy in boosting model performance across diverse domains, including computer vision, natural language processing, security, and healthcare. A comparative analysis shows how these methods improve tasks like image classification, object detection, and medical image segmentation, while also addressing class imbalance and enhancing model robustness on unseen data. Transfer learning is discussed as a complementary strategy to mitigate data scarcity by transferring knowledge between tasks. The authors conclude by suggesting future research directions, such as integrating meta-learning with augmentation and refining GANs to overcome issues like mode collapse, emphasizing that advancements in augmentation techniques will continue to be crucial for addressing data scarcity in deep learning.

3 Methodology

3.1 Data Acquisition and Preprocessing

In this research, a dataset of 120 images, was split evenly across training, validation, and testing sets, with two classes: DR and No DR. This dataset was extracted from a Kaggle dataset[11]. A minimal number of images were picked to test how well models using transfer learning techniques could perform. During preprocessing, operations like rescaling, sample-wise centering, normalization, rotation, width and height shifting, shear transformation, zooming, and flipping were done using ImageDataGenerator[12]. These steps helped train the model effectively, even with the limited dataset.

3.2 Model Selection

The models used for this research are DenseNet201, ResNet152, VGG16, InceptionV3, MobileNet, and EfficientNetB0 which utilize transfer learning and have diverse Architecture with input shape being (224 x224 x3), where 224x224 specifies the height and width of the image and 3 indicates the number of channels for the RGB color (Red, Green, Blue).

DenseNet201[13] is selected due to its dense connectivity, which promotes feature reuse and addresses the vanishing gradient issue. Dense connections between layers allow the model to effectively pass information forward, making it ideal where the dataset is limited, as in this study. Its architecture reduces the number of parameters while ensuring the critical features are captured, an important aspect when working with minimal data.

ResNet152[14] is chosen because of its deep residual architecture, which effectively reduces the vanishing gradient problem. Residual connections allow the network to learn residual mappings, ensuring the gradients flow smoothly even in very deep networks. This architecture has consistently displayed strong performance in various image recognition tasks, making it a reliable choice for this study. Given the complexity of medical images, ResNet152's deep structure ensures capturing complicated features while maintaining accuracy.

VGG16[15] is chosen for its simple and well-established performance in image classification tasks. Its use of small 3x3 convolutional filters enables the model to capture detailed hierarchies in the input images. The relatively small depth of VGG16 (compared to deeper networks like ResNet) makes it suitable for this study, as it reduces the risk of overfitting while still offering sufficient capacity to handle the complexities of retinal fundus images.

InceptionV3[16] is selected for its ability to capture multiscale features using inception modules. This allows the network to process features at different scales simultaneously, making it particularly effective for medical images, where varying levels of detail are important for classification. Additionally, the architecture includes optimizations like factorized convolutions and dimensionality reduction, making it a computationally efficient critical factor given the minimal dataset and resource constraints of this study.

MobileNet[17] is chosen for its efficiency and suitability for low-resource environments. Its use of depthwise separable convolutions reduces the number of parameters and computational costs, making it a good choice for tasks involving limited data. MobileNet's lightweight architecture enables it to generalize well without overfitting, making it particularly effective for transfer learning in resource-constrained settings.

EfficientNetB0[18] is selected due to its balanced approach to scaling network depth, width, and resolution through compound scaling. This allows EfficientNetB0 to achieve high accuracy while remaining computationally efficient. Given the small dataset used in this study, EfficientNetB0's optimized architecture is well-suited for extracting features with minimal computational overhead, making it a strong candidate for diabetic retinopathy detection in this context.

Additionally, batch normalization[12] is employed to enhance the training process of artificial neural networks. This technique standardizes the input of each layer, specifically the activations, by subtracting the batch mean and dividing by the batch standard deviation. This normalization procedure aids in stabilizing and expediting the training process.

Furthermore, Global Average Pooling (GAP)[19] is utilized alongside these pre-trained layers. The GAP layer captures the contextual information of features extracted from the convolutional layers without resorting to fully connected layers, which often contain a large number of parameters. GAP achieves this by averaging each channel of the feature maps, thereby reducing the dimensions of the feature maps to a single value per channel. This reduction in parameters and computational complexity retains all essential information during this step.

At the final layer, a dense layer[12] is incorporated to capture features from preceding layers and classify the data based on the number of classes. Softmax activation[12] is then applied to generate a probability distribution across multiple classes.

3.3 Training

In the study analyzing how different pre-trained models perform, we used the Adam optimizer[20] to train those deep neural networks. Adam's adaptive learning rates and bias correction boost performance compared to previous optimization methods. Alongside Adam, we used the categorical cross-entropy loss function[21]. This function penalizes the model when it assigns low probabilities to the true class by gauging how deviant its predictions are from the actual data distribution. We then use accuracy metrics[22] to see how well the model's performance compares the number of correctly classified instances to the total instances in the dataset.

$$\text{Accuracy} = \frac{Number\ of\ correct\ predicitons}{Total\ Number\ of\ predctions} \times 100\%$$

The training process involves utilizing pre-trained weights from the ImageNet dataset[23], which are frozen to train the model. This freezing is done while applying the fit_generator method[12], training the models for both 5 and 10 epochs.

3.4 Evaluation

In this study, we assessed the trained models on the test dataset using the evaluate_generator function[12], which calculates evaluation metrics like accuracy[22] based on the model's predictions. The test dataset, filled with unseen images, served as a benchmark to check how well the models could predict. We relied on core metrics like accuracy to measure how well the models classified the images. The evaluation process ran through test batches one by one, making sure we got a thorough assessment by setting 'steps=len(test_batches)'. These evaluations gave us insights into the model's ability to classify diabetic retinopathy.

3.5 Visualization

In our research into diabetic retinopathy detection, we utilized Matplotlib, a Python visualization library[24], to analyze the performance metrics of several pre-trained models. Matplotlib creates clear plots that show the key parts of the model training and evaluation, displaying validation accuracy changes over the epochs. We used different colors and line styles for each model's performance trend, making it easy to compare them side by side.

4 Results and Discussion

4.1 Models Trained for 5 epochs

The validation accuracy across different pre-trained models for diabetic retinopathy detection displayed a range of performances. DenseNet201 and VGG16 held steady at around 50% validation accuracy throughout all the epochs. ResNet152 had a variation in its validation accuracy, which ranged between 42.5% and 52.5%, pointing to some inconsistency in how well it performed across the training. On the other side, InceptionV3 showed a steady climb in validation accuracy from 47.5% to 57.5% as training progressed. MobileNet and EfficientNetB0 had validation accuracy ranging from 50% to 55%, which showed decent performance without much improvement over time.

In evaluating the performance of pre-trained models for diabetic retinopathy detection with a minimal dataset and transfer learning, we monitored the validation loss across six CNN models. DenseNet showed promising results, with its validation loss steadily decreasing from 1.4932 to 1.0735 over five epochs, while ResNet had slight fluctuations between 0.7404 and 0.7358. In contrast, VGG16 and InceptionV3 had increased validation losses, with VGG16 rising from 0.8321 to 0.8963 and InceptionV3 going from 0.8818 to 0.9765. MobileNet had a minor decrease, moving from 0.7464 to 0.8267, despite some variability, while EfficientNet maintained a relatively stable validation loss, hovering between 0.7157 and 0.7384. These mixed outcomes underscore how some models improved, while others experienced challenges with fluctuating or increasing validation losses.

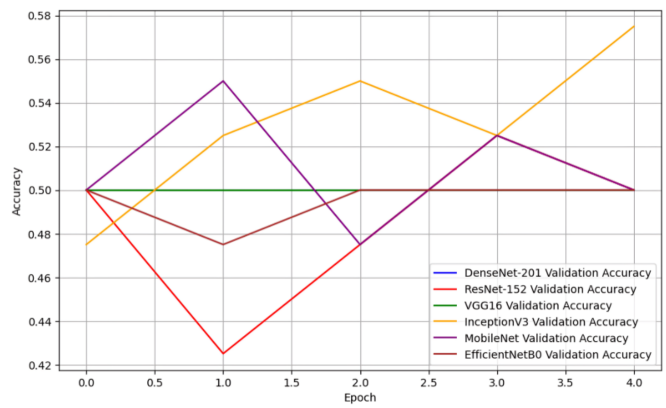

Fig 1. Validation Accuracy of models at 5 epochs

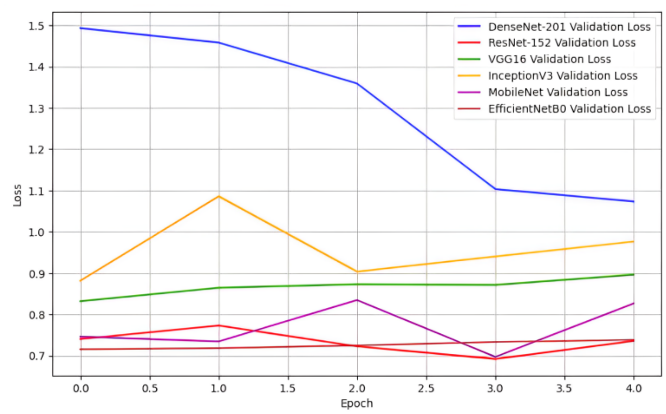

Fig 2. Validation Loss of models at 5 epochs

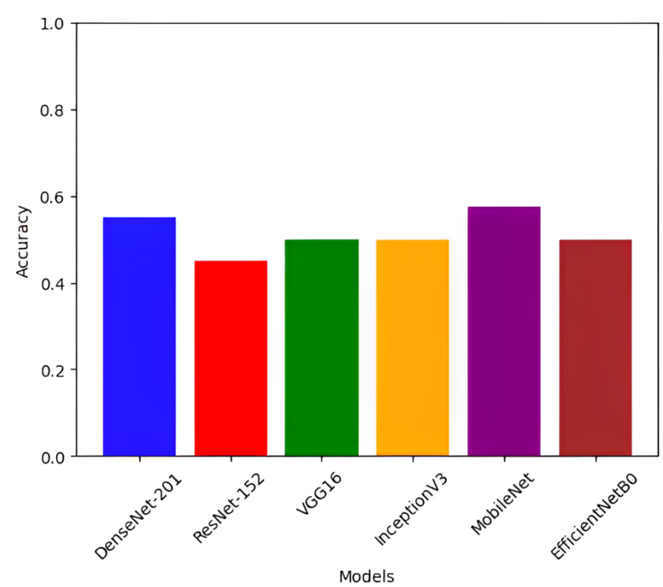

Fig 3. Test Accuracy of models at 5 epochs

In assessing the performance of the models on the test dataset, we observed varying degrees of accuracy. The test accuracy scores were as follows: DenseNet achieved 55.0%, ResNet achieved 45.0%, VGG16 scored 50.0%, InceptionV3 also achieved 50.0%, MobileNet demonstrated the highest accuracy at 57.5%, and EfficientNet matched VGG16 with a score of 50.0%. Notably, MobileNet emerged as the model with the highest accuracy among the evaluated architectures. These findings underscore the importance of model selection in diabetic retinopathy detection, with MobileNet showcasing superior performance in this study.

4.2 Models Trained for 10 epochs

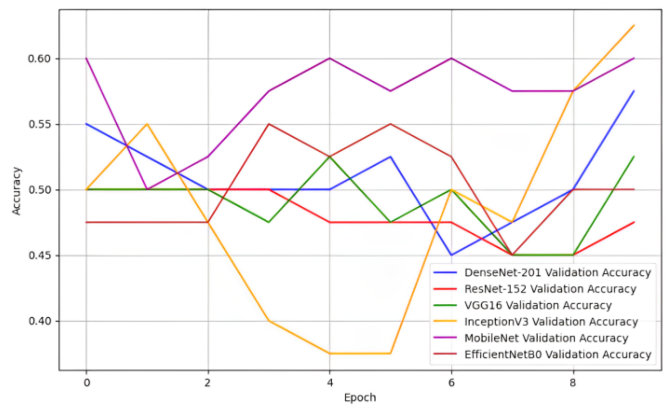

Fig 4. Validation Accuracy of models at 10 epochs

After evaluating the validation accuracy of different CNN models over ten epochs, the results were varied. DenseNet started at 55.0%, got an inclination to 57.5%, and has been stable until it ended at 57.5%. ResNet, on the other hand, had a constant 50.0% accuracy all the way through. VGG16 showed some ups and downs, starting at 50.0%, dropping to 47.5%, and then bouncing back to 52.5% by the end. InceptionV3 had a peak at 62.5% but then slid down to 57.5%. MobileNet was stable, beginning at 60.0% and staying around that mark for most of the epochs. EfficientNet started at 47.5%, hitting a high of 55.0%, and finished at 50.0%. In the end, MobileNet was the standout, showing the highest and most consistent validation accuracy among the models.

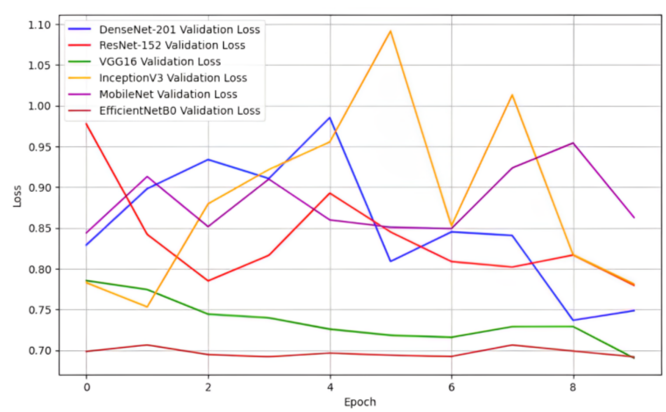

Fig 5. Validation Loss of models at 10 epochs

Assessing the validation loss trends across ten epochs for the different models, DenseNet started with a validation loss of 0.8291, then decreased to 0.7484, getting better over time. ResNet had a similar downward trend, going from 0.9774 to 0.7795. VGG16, on the other hand, is unstable, with its validation loss bouncing between 0.7258 and 0.7853. Inception V3 also had its ups and downs, with loss numbers swinging from 0.7827 to 0.7812. MobileNet stayed steady, with its validation loss between 0.8439 and 0.8629. EfficientNet is consistent, sticking close to 0.6946. Overall, these trends show how different the models performed, with some showing solid improvements while others had more stable or unpredictable validation losses.

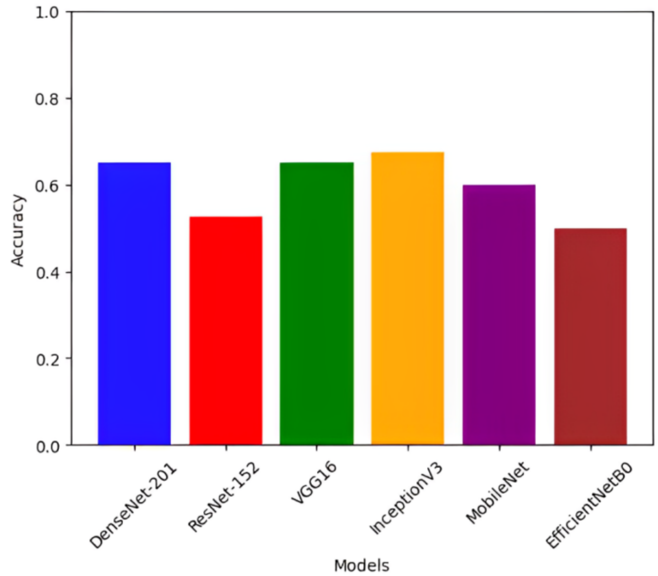

Fig 6. Test Accuracy of models at 10 epochs

The test accuracy scores for the evaluated models reveal varying degrees of performance. DenseNet achieved an accuracy of 50.0%, Similarly, VGG16 demonstrated a test accuracy of 52.0%, and InceptionV3 achieved 55.0%. Notably, MobileNet emerged as the top performer with a test accuracy of 60.0%. In contrast, ResNet and EfficientNetB0 exhibited the lowest accuracy among the evaluated models, standing at 42.5%. These findings underscore the importance of model selection and optimization in achieving superior performance in diabetic retinopathy detection tasks.

4.3 Models Trained for 15 epochs

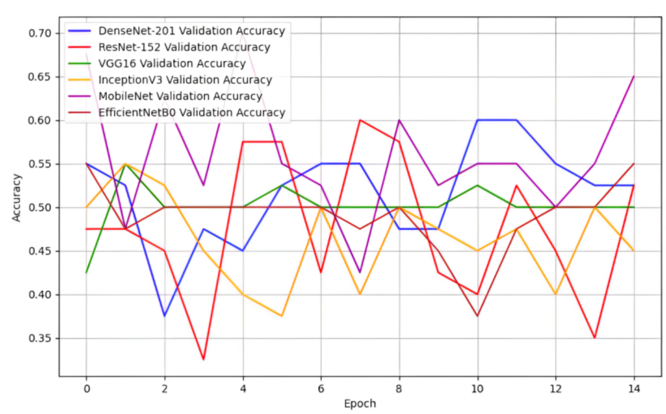

Fig 7. Validation Accuracy of models at 15 epochs

The validation accuracy scores for the evaluated models varied across the different architectures. For the DenseNet model, the validation accuracy ranged from 37.5% to 60.0% across the 15 epochs, with fluctuations throughout training. The ResNet model exhibited validation accuracies between 32.5% and 60.0%, showing similar fluctuations during training. VGG16 also displayed fluctuating validation accuracies, ranging from 42.5% to 70.0% over the 15 epochs. Inception's validation accuracies varied between 40.0% and 55.0%, showcasing a relatively stable but modest performance. MobileNet demonstrated validation accuracies between 47.5% and 70.0%, indicating better performance compared to some other architectures. Lastly, EfficientNet showcased validation accuracies ranging from 45.0% to 55.0%, with fluctuations during training.

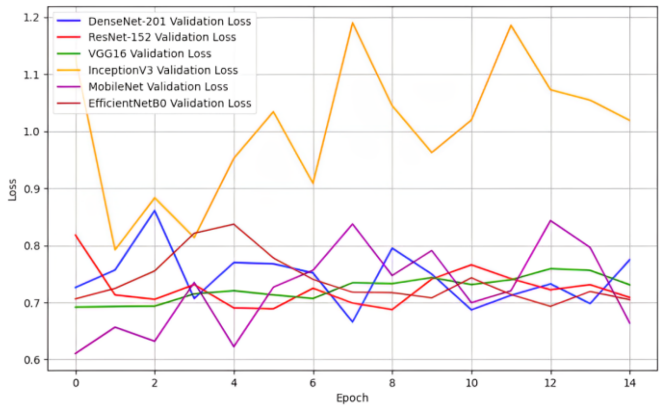

Fig 8. Validation Loss of models at 15 epochs

The validation loss during the training of different models varied across the epochs. For DenseNet, the validation loss ranged from 0.7265 to 0.7749, with fluctuations observed throughout the epochs. ResNet exhibited a similar trend, with validation loss fluctuating between 0.7133 and 0.8180. VGG16 demonstrated relatively stable validation loss values, hovering around 0.6919 to 0.7436. Inception, on the other hand, showed more erratic behavior, with validation loss varying between 0.7922 and 1.1903. MobileNet displayed a decreasing trend in validation loss from 0.6107 to 0.6641. Finally, EfficientNet exhibited fluctuating validation loss values, ranging from 0.6932 to 0.8371.

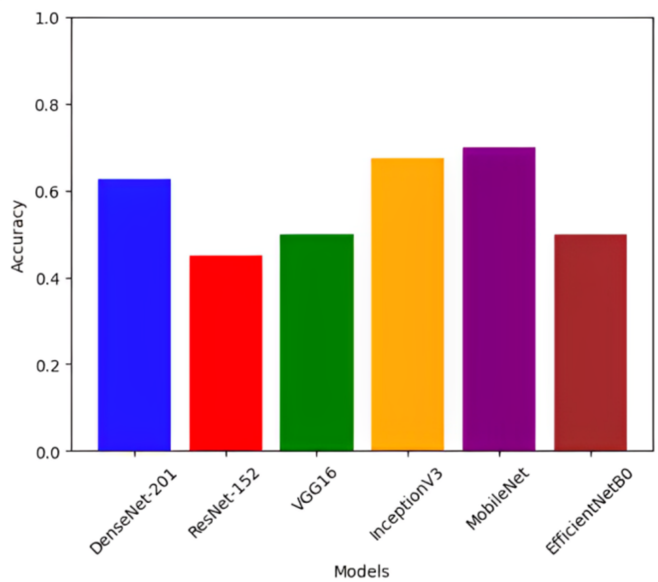

Fig 9. Test Accuracy of models at 15 epochs

The test accuracy scores for various models were evaluated in the study. DenseNet201 achieved an accuracy of 62.5%, ResNet152 reached 45.0%, VGG16 attained 50.0%, InceptionV3 reached 67.5%, MobileNet performed the best with an accuracy of 70.0%, and EfficientNetB0 achieved 50.0%. These models were assessed for their effectiveness in detecting diabetic retinopathy. Notably, MobileNet emerged as the top performer among them, demonstrating the highest accuracy on the test dataset. These findings underscore the importance of selecting appropriate models for diabetic retinopathy detection, with MobileNet showing promising results in this context.

Comparing the accuracy of models trained on a minimal dataset in this study against those from previous research [25], some noticeable differences have been noticed. At 15 epochs, MobileNet achieved 70.0% accuracy, followed by InceptionV3 at 67.5%, DenseNet201 at 62.5%, and VGG16 in the end at 50.0%. On the other side, past studies have shown much higher accuracy for the same models trained on larger datasets. For example, InceptionV3 and ResNet50 clocked in at 89.7% and 87.1%, respectively. The lower performance here likely stems from the limited dataset, which hampered the model's ability to generalize effectively. Additionally, models like EfficientNetB5 had accuracy levels up to 92.52% in other studies; our run with EfficientNetB0 got up to 50.0%, showcasing the importance of data quantity and diversity for model performance. These results bring out the challenges and limitations of using transfer learning when dealing with minimal datasets, stressing the importance of careful model selection and optimization.

In contrast to previous studies which have utilized larger and more varied datasets for detecting diabetic retinopathy, this study dives into the challenges and potential of using transfer learning on minimal datasets. Models like MobileNet and DenseNet201 held their own, considering the limited data, but the results highlight how crucial dataset size and diversity are for top-tier model performance. What's new about this research is its focus on tweaking pre-trained models to work in low-resource settings where data is scarce. While previous studies were all about maxing out accuracy with big datasets, this study sheds light on the trade-offs and careful considerations needed when working with limited data. The findings show that even though the models didn't showcase the high accuracy levels seen in other studies, they still offer a solid approach for diabetic retinopathy detection under specific conditions. Future research could level up by trying out data augmentation or fine-tuning strategies to push model performance even further in similar settings. This study highlights the understanding of how CNN models stack up when data is restricted, laying down a foundation for future work in low-data environments.

5 Conclusion

In this study, we have tested pre-trained CNN models for diabetic retinopathy detection using a very limited dataset of just 40 images for training, 40 for validation, and 40 for testing. Despite these constraints, MobileNet consistently showed top accuracy at 5, 10, and 15 epochs. While InceptionV3 had a strong start, MobileNet showed stable results, proving its adaptability and robustness with minimal data. This study's novelty lies in its focus on transfer learning functionality in resource-constrained settings, an area unexplored in existing literature. For future work, experimenting with data augmentation and fine-tuning strategies could boost model performance even further. Expanding the dataset and testing newer models could also provide more insights. Conclusively, MobileNet stands out as a reliable option for detecting diabetic retinopathy in low-resource settings, showing that AI-driven diagnostics can still be effective with minimal data.

5.1. Limitations

The primary limitation of this study is the size and diversity of the dataset. With only 120 images, split evenly between training, validation, and testing, the dataset is significantly smaller than the datasets used for deep learning models, which can range into thousands. This limited data size restricts the models' ability to generalize to new, unseen data, as seen by the lower accuracy rates compared to previous studies that used larger datasets.

Additionally, a lack of diversity in the dataset. Diabetic retinopathy images can vary based on factors such as lighting conditions, and camera resolution. As a result, models trained on this dataset may perform well on similar, controlled images but struggle when exposed to real-world data with higher variability.

The minimal dataset also increases the risk of overfitting, where models learn the specific features of the small dataset rather than general patterns that can apply to broader data. Although transfer learning helps mitigate this risk by using pre-trained models, the potential for overfitting remains.

References

1) Shukla UV, Tripathy K. Diabetic Retinopathy. 2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK56080/.
2) Azagew AW, Yohanes YB, Beko ZW, Ferede YM, Mekonnen CK. Determinants of diabetic retinopathy in Ethiopia: A systematic review and meta-analysis. . *PloS One*. 2023;18. Available from: https://doi.org/10.1371/journal.pone.0286627.
3) Alam S, Hasan MK, Neaz S, Hussain N, Hossain MF, Rahman T. Risk Factors, Diagnosis, Complications and Comprehensive Management. *Diabetes Mellitus: Insights from Epidemiology*. 2021;2(2):36–50. Available from: https://doi.org/10.3390/diabetology2020004.
4) Pm E, Ranjana P. An Approach Using Transfer Learning to Disclose Diabetic Retinopathy in Early Stage. *International Conference on Futuristic Technologies in Control Systems & Renewable Energy (ICFCR) 2020*. 2020. Available from: https://doi.org/10.1109/icfcr50903.2020.9.
5) R RR. An Economic Analysis for the Use of Artificial Intelligence in Screening for Diabetic Retinopathy in Trinidad and Tobago. *Cureus*. 2024;16(3). Available from: https://doi.org/10.7759/cureus.55745.
6) Kandel M, Castelli M. Transfer Learning with Convolutional Neural Networks for Diabetic Retinopathy Image Classification: A Review. *Appl Sci*. 2020;10(6):2021–2021. Available from: https://doi.org/10.3390/app10062021.
7) Salehi AW. A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges Future Scope. *Sustainability*. 2023;15(7):5930–5930. Available from: https://doi.org/10.3390/su15075930.

8) Tymchenko B, Marchenko P, Spodarets D. Deep learning approach to diabetic retinopathy detection. *arXiv preprint* . 2020. Available from: https://doi.org/10.48550/arXiv.2003.02261.
9) Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: a literature review. *BMC medical imaging*. 2022;22(1):69–69. Available from: https://doi.org/10.1186/s12880-022-00793-7.
10) Bansal MA, Sharma DR, Kathuria DM. A systematic review on data scarcity problem in deep learning: solution and applications. *ACM Computing Surveys (Csur)*. 2022;54:1–29. Available from: https://doi.org/10.1145/3502287.
11) Dola1507108. Diabetic Retinopathy Classified [Data set]. 2024. Available from: https://www.kaggle.com/datasets/dola1507108/diabetic-retinopathy-classified.
12) Wang K, Chen C, He Y. Research on pig face recognition model based on keras convolutional neural network. *InIOP Conference series: earth and environmental science* . 2020;474(3):32030–32030. Available from: https;//doi.org/10.1088/1755-1315/474/3/032030.
13) Pawar SD, Sharma KK, Sapate SG, Yadav GY, Alroobaea R, Alzahrani SM, et al. Multichannel DenseNet architecture for classification of mammographic breast density for breast cancer detection. *Front Public Health*. 2022;10:885212–885212. Available from: https://doi.org/10.3389/fpubh.2022.885212.
14) Shafiq M, Gu Z. Deep Residual Learning for Image Recognition: A Survey. *Applied Sciences*. 2022;12(18):8972–8972. Available from: https://doi.org/10.3390/app12188972.
15) Tammina S. Transfer learning using vgg-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications (IJSRP)*. 2019;9(10). Available from: https;//doi.org/10.29322/IJSRP.9.10.2019.p9420.
16) Pan Y, Liu J, Cai Y, Yang X, Zhang Z, Long H, et al. Fundus image classification using Inception V3 and ResNet-50 for the early diagnostics of fundus diseases. *Frontiers in Physiology*. 2023;14. Available from: https;//doi.org/10.3389/fphys.2023.1126780.
17) Wang W, Li Y, Zou T, Wang X, You J, Luo Y. A novel image classification approach via dense-MobileNet models. *Mobile Information Systems*. 2020;2020:7602384–7602384. Available from: https://doi.org/10.1155/2020/7602384.
18) Shah HA, Saeed F, Yun S, Park JH, Paul A, Kang JM. A robust approach for brain tumor detection in magnetic resonance images using finetuned efficientnet. *IEEE Access*. 2022;10:65426–65464. Available from: https://doi.org/10.1109/ACCESS.2022.3184113.
19) Akhtar N, Ragavendran U. Interpretation of intelligence in CNN-pooling processes: a methodological survey. . *Neural computing and applications*. 2020;32:879–98. Available from: https://doi.org/10.1007/s00521-019-04296-5.
20) Gaddam DK, Ansari MD, Vuppala S, Gunjan VK, Sati MM. A performance comparison of optimization algorithms on a generated dataset. In: InICDSMLA 2020: Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications 2022. Springer. ;p. 1407–1415. Available from: https://doi.org/10.1007/978-981-16-3690-5_135.
21) Ciampiconi L, Elwood A, Leonardi M, Mohamed A, Rozza A. A survey and taxonomy of loss functions in machine learning. *arXiv preprint* . 2023. Available from: https://doi.org/10.48550/arXiv.2301.05579.
22) Naidu G, Zuva T, Sibanda EM. A Review of Evaluation Metrics in Machine Learning Algorithms. In: Lecture Notes in Networks and Systems. Springer International Publishing. 2023;p. 15–25. Available from: https://doi.org/10.1007/978-3-031-35314-7_2.
23) Shankar V, Roelofs R, Mania H, Fang A, Recht B, Schmidt L. Evaluating machine accuracy on imagenet. *InInternational Conference on Machine Learning*. 2020;p. 8634–8644. Available from: https://proceedings.mlr.press/v119/shankar20c.html.
24) Lemenkova P. Python libraries matplotlib, seaborn and pandas for visualization geo-spatial datasets generated by QGIS. . *Analele stiintifice ale Universitatii*. 2020;64:13–32. Available from: https://doi.org/10.6084/m9.figshare.13010069.
25) Sharma N, Mangla M, Iqbal MM, Mohanty SN. Deep Learning Framework for Identification of Skin Lesions. . *EAI Endorsed Trans Perv Health Tech*. 2019;9. Available from: https://publications.eai.eu/index.php/phat/article/view/3900.