# Deep Learning-Based Automated Detection and Classification of Diabetic Retinopathy Using MobileNetV2 and DenseNet201

MSc Research Project
Artificial Intelligence 2024 - 2025


Aswin Kumar G R
Student ID: x23245778

School of Computing
National College of Ireland


Supervisor: Ms. Sheresh Zahoor

**National College of Ireland**
**Project Submission Sheet**
**School of Computing**

| | |
|---|---|
| **Student Name:** | Aswin Kumar G R |
| **Student ID:** | x23245778 |
| **Programme:** | Artificial Intelligence |
| **Year:** | 2024 - 2025 |
| **Module:** | MSc Research Project |
| **Supervisor:** | Ms. Sheresh Zahoor |
| **Submission Due Date:** | 12/12/2024 |
| **Project Title:** | Deep Learning-Based Automated Detection and Classification of Diabetic Retinopathy Using MobileNetV2 and DenseNet201 |
| **Word Count:** | 7891 |
| **Page Count:** | 23 |

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

**ALL** internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

| | |
|---|---|
| **Signature:** | Aswin Kumar G R |
| **Date:** | 4th December 2024 |

## PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

| | |
|---|---|
| Attach a completed copy of this sheet to each project (including multiple copies). | ☐ |
| **Attach a Moodle submission receipt of the online project submission**, to each project (including multiple copies). | ☐ |
| **You must ensure that you retain a HARD COPY of the project**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. | ☐ |

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

| **Office Use Only** | |
|---|---|
| Signature: | |
| Date: | |
| Penalty Applied (if applicable): | |

# Deep Learning-Based Automated Detection and Classification of Diabetic Retinopathy Using MobileNetV2 and DenseNet201

Aswin Kumar G R
x23245778

**Abstract**

Diabetic Retinopathy (DR) is one of the most common causes of blindness in global health today, early screening is important in preventing the diseases progression and complications. The existing approaches in diagnosing TB are time-consuming, involves the use of expert personnel which is barely available for timely diagnostic intervention. Although the automated solutions based on deep learning have been demonstrated as feasible, a range of unvarying and efficient classification across all the levels of DR stage has not been achieved. This study leverages state-of-the-art deep learning models, MobileNetV2 and DenseNet201, to classify retinal images from the Kaggle Retinopathy dataset into five DR stages: It can be None, Mild, Moderate, Severe and Proliferative Diabetic Retinopathy abbreviated as No_DR, Mild, Moderate, Severe, Proliferative_DR. Normalization was used with the aims to make the model more launches and insensitive to the input data values, as well as, data augmentation was performed to minimize class over/underrepresentation. Fine-tuning these models was done while employing common metrics, including accuracy rate, precision, recall rate, and F1-coefficient. The results therefore validate these models, with MobileNetV2 attaining 93.4% accuracy and F1-score, while DenseNet201 attained 92.5% accuracy and F1-score. In view of these, it is evidenced that combining efficient deep learning techniques for scalable, accurate, and efficient DR classification allows enhanced diagnostic reliability in clinical practice.

## 1 Introduction

Diabetic Retinopathy (DR) is an important complication of diabetes and the major cause of vision loss and blindness among working-aged adults. Recently, the global prevalence reached 463 million and is expected to increase DR burden significantly International Diabetes Federation (2019). Diabetic retinopathy (DR) is characterised by morphological changes and classified by the severity of DR into non-proliferative DR and proliferative DR which often leads to blindness. Prevention is the best cure, especially in Kadima where early detection of these diseases received very little support, despite the fact that treatments for blinding diseases are available; the current techniques are however very costly in term of manpower and time as they depend on a team of ophthalmologists to analyze the images of the retina. The inequality of global healthcare leads to significant challenges particularly in developing and pointedly resource-scarce environment conveys that millions endure unacknowledged and untreated. Alternatively, an automated approach to DR stage detection and classification is feasible. The incorporation of Artificial Intelligence (AI) and Deep Learning especially Convolutional Neural Networks (CNNs) has enabled medical imaging find tools that help with pattern recognition in imaging data types. CNNs have proven to have great promise in areas such as oncology, and radiology in particular where analysis of images enables the determination of disease and facilitates crafting a treatment plan. However, it must be noted that current development of DR classification has open issues such as the variation in image quality, the disparities of the imaging devices, existence of the imbalanced classes of the datasets, and dissimilarities in the performances of automated models Abràmoff and et al. (2018) Solving these issues is crucial for developing valid AI-based diagnostic instruments applicable to clinical practice.

Current DR detection procedures involve the assessment of fundus images by human raters, specifically ophthalmologists, this process is subjective and inter-observer inconsistency. These problems can be addressed by the automated system for DR detection where the classification method of retinal images into DR stages is objective, reproducible, and efficient. Several algorithms have been proposed for this purpose; however, most of them use a combination of hand crafted features or low level vision techniques which may not be efficient for different datasets. Medical image classification has shifted to deep learning because it automatically extracts hierarchical features from the data automatically. However, for the majority of deep learning models for DR detection, the level of accuracy is insufficient or there are no adequate resistance for using in clinic. This is mainly because many issues arise from the handling of imbalanced datasets and the ability to generalize across a range of imaging conditions.

To fill these gaps, this research studies the performance of two state of the art CNN architectures; MobileNetV2 and DenseNet201 for DR stages classification as aided by the Kaggle Retinopathy dataset. MobileNetV2 is considered light-weighted, thus best suited for deployment in environments with limited resources while DenseNet201 is well suited for feature extraction because of the densely connected layers which make reusing features easier. By leveraging transfer learning, this study fine-tunes these pre-trained models to classify retinal images into five DR stages: No DR, Mild, Moderate, Severe, Proliferative DR. Pre-trained layers help in reducing the amount of labeled datasets needed and on top of it the robust layer enables the models to learn from the large datasets it has been trained from like the ImageNet Deng and et al. (2009). In any deep learning, the preprocessing phase is important, especially if the data set in the medical images is imbalanced. As part of the preparatory measures in this study, a significant number of steps were taken to improve model performance and versatility. Data normalization techniques applied in the set pre-processing floor made pixel intensity distributions standardized, while introduction of data augmentation introduced some form of variation in the training data, having a positive impact on the prevention of overfitting. These steps were necessary in dealing with issues that characterizes Retinopathy dataset including differences in images sizes, illumination and class imbalance problem.

## 1.1 Research Question

The primary research question guiding this work is: ***How do MobileNetV2 and DenseNet201 models perform in categorizing retinal images from the Dia-***

***betic Retinopathy Detection dataset into the 5 stages of Diabetic Retinopathy in terms of accuracy, precision, recall, and F1-score?***

To address this question, the following objectives were outlined:

1. Review the current literature on machine learning algorithms for automatic DR diagnosis.

2. Develop transfer learning models, including MobileNetV2 and DenseNet201, for multi-class classification.

3. Employ technical pre-processing measures adaptable to the features of the analysed databases.

4. Assess the performance of the presented approaches in terms of clinical relevance. By systematically addressing these objectives, this work aims at providing the proof-of-concept of the utilization of advanced CNNs for scalable DR classification.

The findings of this study offer important information to understand the application of deep learning in DR detection in detail. In all the stages of DR, MobileNetV2 model was as accurate and reliable as DenseNet201 model. MobileNetV2 achieved 93.4% of accuracy and F1 score of 93.4%, where DenseNet201 achieved an accuracy of 92.5% and F1 score of 92.5%. These metrics provide evidence for the performance of such models in capturing patterns in images of retinas while providing accurate classification even where these patterns may be intricate. Moreover, the models developed in this paper incorporate transfer learning and rigorous data preprocessing to guarantee that the model's performance is not only high but also stable in terms of input variations. These observations open up the possibility of these architectures for narrowing the gap between conceptional models of AI and actual usable solutions.

In theoretical framework of the work, it inherits to the development of recognizing deep learning in medical imaging by showcasing the use of improved CNN frameworks for multi-classification of the DR stages. The incorporation of MobileNetV2 and DenseNet201 to the diagnostic process has further possibilities to improve the DR screening portability and availability in the developing countries. However, in the real-word application, such high accuracy and reliability of these models, may imply shorter diagnostic time or better patient's prognosis, not to mention reduced healthcare costs. Overall, the automatic classification leaves clinicians to concentrate on the treatment and the patients while using the outcome as an input from AI. Nevertheless, it is essential to note the following objectives of this study. The Retinopathy dataset, despite being extensive, may not be sufficient to depict all the varying state of real-world retinal images. In the same regard, transfer learning implies that the weights they incorporate are initially trained on other tasks as well, and hence the models have access to limited optimized knowledge for medical imaging. In future studies, the information could be gathered from larger and more divers samples, and the methods of training the models for different domains could be stimulated for better results.

Finally, this study aims to show that both MobileNetV2 and DenseNet201 can be powerful instruments for the automated classification of DR. In gaining high accuracy and precision with high recall as well as high F1-score, it indicates that these models can be applied in different clinical applications since both scalability and reliability are essential in clinical engineering solutions. Besides contributing to the state of the art in DR detection, this paper also provides a foundational research for the next wave of improvement and development of AI-based diagnostic systems in the medical filed.

The structure of this paper is as follows: In the Literature Review section reviews previous attempts at DR identification and the challenges of those methods to situate this work. Section 3 describes the data set, data pre-processing and training of MobileNetV2 and DenseNet201 networks. The Results and Discussion sections describe the evaluation indicators of the constructed models and discuss the potential applications in clinical practice. Lastly, the Conclusion restates the findings in light of the subject and contributes positively toward the area of medical diagnostics.

# 2 Literature Review

Diabetic Retinopathy(DR) is a common microvascular complication of diabetes and a leading cause of blindness among the populace worldwide. As evidenced in this paper, DR detection has become an increasingly complex problem as new AI methods such as ML and DL emerge; many prior and recent studies have harnessed these techniques to build automated DR detection systems. This section provides a comprehensive review of related work in the domain, organized into three sections: classification methods using early traditional approaches, deep learning improvements, current trends and issues. Included in the review are the datasets employed, the models utilized, the accuracy obtained, and the constraints of each type of strategy, followed by a discussion of how the current methods fall short and where and how they could be improved.

## 2.1 Conventional Machine Learning Techniques

In early endeavors of DR detection, strong reliance was put on conventional machine learning methods with respect to engineered features. These models normally employed low level image features including color, texture and shape of the retinal image to detect DR related pathological features including microaneurysms, exudates and haemorrhages. Niemeijer and et al. (2007) proposed one of the first algorithms for identifying microaneurysm, using DiaretDB1 dataset. By using the k-Nearest Neighbors (k-NN) classifier, the authors obtained the classification accuracy of 83%. While engendering great results, such a model was criticized for a high false positive rate stemming from the fact that microaneurysms often resemble normal retinal structures. In turn Adal and et al. (2010) improved the field by employing a Support Vector Machine (SVM) classifier for red lesion detection in images of the retina. Their study used the Messidor dataset and was 86% accurate. SVMs overall provided high accuracy but were rigid to the data set while also requiring time-consuming feature engineering.Such classical methods served as the basis of automated DR diagnosis. However, they introduced critical challenges; those designs required domain-specific knowledge to build features, and limited scalability could not scale the large dataset with a variation of picture expertise.

## 2.2 Advances with Deep Learning

Diabetic retinopathy is one of the complications of diabetes which if not diagnosed will result in blindness. Alyoubi et al. (2023) surveyed deep learning approach focusing on CNNs for DR detection , where they discussed the medical image analysis and pointed to the challenges for future studies. Similarly, Rajalakshmi et al. (2023) evaluated an AI-based software with a fundus photograph taken by a Smartphone and showed sensitivity of 95.8% and specificity of 80.2% for DR and proposed its usefulness for a mass examination. In a similar study, Bhandari et al. (2023) worked on detecting new onset DR and grading its severity at the earlier stage with help of some soft computing approaches: Particle Swarm Optimization (PSO), Genetic Algorithm (GA), and Neural Networks, deep learning. Gangwar and Ravi (2023) developed the hybrid deep learning model by using transfer learning on Inception-ResNet-v2; the proposed model tends to provide the higher accuracy of 72.33% on Messidor-1 and 82.18% on APTOS. New architectures, including the VGG-NIN model proposed by Khan et al. (2023) built based on the combination with VGG16, SPP, and NiN layers to boost nonlinear feature extraction, demonstrated better performance and computational reconstruction required in DR classification. In combination, these studies demonstrate that DR is best detected and graded using advanced machine learning and deep learning approaches at scale and with high accuracy.The use of deep learning especially signaled a shift in how DR was detected directly from raw images given that the approach permitted automated feature extraction and analysis. A new and powerful type of DL approach known as the Convolutional neural networks (CNNs) became the workhouse of most of the state-of-the-art systems and greatly enhanced both the detection accuracy and robustness of the systems. Gulshan and et al. (2016) the authors were one of the pioneers who used a CNN for DR detection. By employing the EyePACS dataset here they proposed a deep CNN for operating the categorization of fundus images with respect to DR or non DR condition. Through more detailed model, they got the sensitivity of 97.5%, specificity of 93.4% and it proved to be more effective than previous methods. However, they found that while the model had high performance, it required high computational resources and that the model was black box which led to issues with interpretability and access for people with disabilities.A new discovery came from Abramoff and et al. (2018) who proposed a fully-self sustaining AI system for DR detection.

## 2.3 Challenges in Diabetic Retinopathy Detection Using CNNs and Transfer Learning

By integrating the CNNs into their framework and using their own dataset, the authors realized the AUC of 0.98 for the receiver operating characteristic curve. For clinical images, the performance of the system was outstanding except for the fact that proper acquisition of high-quality fundus images cannot always be guaranteed in such environments or Centres with inadequate resources. Some other scholars studied transfer learning to take advantage of prior models. Takahashi and et al. (2017) used the prebuilt image as VGG16 and ResNet50's Messidor and Kaggle to detect DR. On the Messidor dataset, these models were accurate nearly 91 percent thus proving the concept of transfer learning. However, getting carried away with colourful descriptors and getting overly obsessed with achieving extremely high levels of accuracy on these kinds of smaller datasets and the computational challenges involved still remained its drawbacks. Some of the research papers focused on ensemble techniques where the DR detection was found to be more robust. Gargeya and Leng (2017) proposed an ensemble of multiple CNNs, for this study they obtained 0.94 au c on the EyePACS database. While enhancing this structure's detection performance, the findings imply increased computational conditions and training time. Later, these attention mechanisms were incorporated into the CNNs in order to attend to the most relevant regions of the fundus images.

Liu and et al. (2019) implemented an attention-based CNN model for the identification of the IDRiD dataset where model achieved the accuracy of about 89%. Through increasing the understanding of these models, the level of confidence in employing AI for decision making was increased. However, training attention mechanisms demanded more (larger) datasets and extra processing power. The multi-class classification of the DR severity stages was also an active area of focus. MobileNetV2 a light CNN architecture with its inception has been used by Kumar and et al. (2020) to stage DR based on the APTOS Demo dataset. Despite the fact that the model succeeded in achieving an 88 percent accuracy rate, it had some limitations because of the lack of ability to transfer across one set of data to another set of data. In the same way, Wang and et al. (2020) used ResNet scheme to classify DR severity using the Messidor dataset and achieved an accuracy level of 90%. Regarding the intermediate stages of DR more difficultly was observed to differentiate between them because of the similarities in their characteristics. Methods of preprocessing and data augmentation were described to cope with the problem of imbalanced datasets. Garcia and et al. (2021) used histogram equalization and other numerous data augmentation methods to enhance the accuracy of Kaggle dataset containing CNN to 87%. However, these preprocessing steps made the pipeline longer and slower, in terms of the time needed to execution this pipeline. Finally, another developed work was the GANs for synthesising more images of the retinal databases. Chakrabarti and et al. (2021) revealed that the use of generated images by GAN produced a significant enhancement of the model robustness with a corresponding increase of classification by 5%. Nevertheless, training GANs posed problems by consuming vast volumes of computational assets, hence restricting the MNLI's real-world application.

## 2.4 Recent Trends and Challenges

New trends in DR detection have shifted to Light models, Explainable AI and multimodal which seem to solve some drawbacks of traditional and deep learning methods. The proposed architectures like EfficientNet-B0 and MobileNetV2 are designed in such a way so as to enable real time DR screening even at the remote locations where computational power is a constraint. Authors Yuan and et al. (2022) used EfficientNet-B0 for detecting DR in the APTOS dataset with an accuracy of 85 percent. The time optimization was achieved flexibly at the expense of efficiency in handling complex DR stages on the model which was designed for mobile devices. The same way, Hossain and et al. (2023) protype CNN for real-time DR detection with accuracy of 87% of on Messidor dataset. But the results showed that the model worked worse if it got low-quality images as an input. This paper discusses the area of explainability to help improve the accountability and reliability of Artificial Intelligence. Ribeiro and et al. (2021) used saliency maps and Grad-CAM to the fundus images and integrated these into the proposed model, ResNet50. Their model was accurate at 88% on the Kaggle dataset but used extra computational power to produce explanations. In their study, Ahmad and et al. (2022) applied SHAP (SHapley

Additive exPlanations) values to interpret the predictions of their MobileNetV2 model with a study on the EyePACS using accuracy of 93%. For the trade-off of enhancing the user trust, the performance of SHAP values compromised the time for prediction. Multimodal data fusion has also being proved to enhance DR detection results. Zhou and et al. (2021) fused the fundus images with clinical structured data applying the CNN and gradient boosting method and secured 92% accuracy on the EyePACS dataset. Implemented here, this helped to improve the computational decision making in the model but also was dependent on the availability of EHRs. Further development of a multimodal approach was made by Yang and et al. (2022), who included fundus images with optical coherence tomography (OCT). Their model had AUC of 0.95 but it need costly, OCT instrument that is impractical in a low income country. ViT, as one of the recent architectures, has gained remarkable outcomes in the DR detection field. These authors used ViTs on the IDRiD dataset, where the accuracy was found to be 0.96, as measured from the AUC. Categorically, the proposed method, the ViTs, provided relatively higher levels of accuracy and interpretability than the CNNs while suffering from high computational costs. However, there are still some issues that can be considered to be critical for improving and further developing reliable, scalable, and understandable models. The need for large annotated training sets and large amounts of labeled data remain challenges along with the sensitivity of the method to image quality and the black box nature of deep learning models. In addition, the application of the advanced architectures has the disadvantage of requiring massive computations an aspect that is unattainable with limited resources.

From the reviewed works, the authors unbiasedly shown that great advancement has been achieved towards diagnosing diabetic retinopathy through the use of traditional and deep learning methods. The classical approaches of learning models were good but not learnable and not very robust and also not scaleable. CNNs were the main deep learning models that helped change the DR detection by removing the need for feature extraction by the human eye and increasing accuracy. Some of the earlier methods' shortcomings have been offset by the innovation of lightweight architectures, attentions schemes, and multimodality. However, some issues still remain, such as the requirement of a big amount of annotated images, the problem of Model interpretability, and their sensitivity to image quality. Missing from the state-of-the-art are models that are both lightweight and easy to interpret and specifically designed for deployment in areas of low computational resources. On the basis of such limitations this research proposes to fill these gaps by creating a comprehensive, explainable, and fast model for DR identification.

# 3 Research Methodology Design Implementation

Diabetic retinopathy (DR) would be amongst the primary reasons for blindness and vision loss for people with such conditions as diabetes. Consequently, timely identification and correct stratification of the patients according to the severity of DR is paramount in management and obliteration of vision loss. In this study, DenseNet201 and MobileNetV2 machine learning models are used as a means of detecting and differentiating DR with the use of retinal imagesCreswell (2014). These two models are considered more efficient in image classification tasks than many of the most advanced models; both of them show the ability to perform transfer learning using pre-trained ImageNet weights; and both are known to work quite well with relatively minimal training when applied to domain-specific data sets. Generally, DenseNet201 uses densely connected convolutional network to adapt grad-centre propagation and feature repetition, while MobileNetV2 is lightweight and fast to use and perfectly suited in real-time application in the limited resources environment. Methodology of this study involves, data gathering/collection and cleaning as well as preparation of data, model training, model assessment, and implementation as shown in Figure 1. This section also includes visual figures and tools that support and explain the processes and outcomes arrived at.

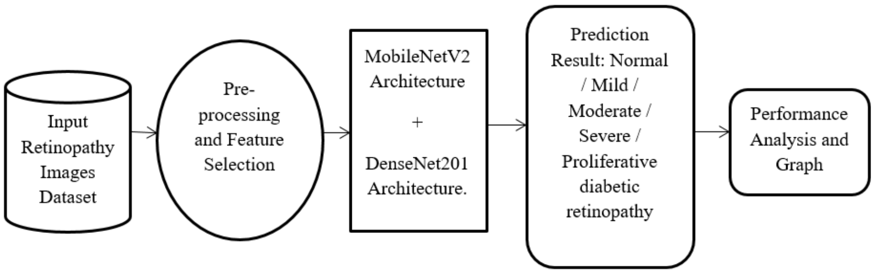


Figure 1: Model Workflow

## 3.1 Model Selection

Thus, DenseNet201 and MobileNetV2 were selected as the main models for this study the reason being that they offer the best results when it comes to image classification. (DenseNet201) is a densely connected convolutional neural network in which every node is connected to all previous nodes which was proposed by Huang, Liu, Maaten and Weinberger (2017). This architecture enables reuse of the features, solving the vanishing gradient problem, and proper use of parameters. These attributes make DenseNet201 best suited to learn fine-level features from high-resolution retinal images including blood vessels and microaneurysms that indicate different levels of DR.

The second model we tested is MobileNetV2 proposed by Sandler et al. (2018), which is designed for mobile and other edge devices. Its architecture uses inverted residual blocks and depthwise separable convolutions; this has reduced computational work by a large extent but does not compromise with the accuracy achieved. The efficiency of MobileNetV2 means that it is possible to deploy it for real-time use – as a component of an online diagnostic tool. The necessary balance of DenseNet201 and MobileNetV2 guarantee that the models reflect both the high accuracy and the high efficiency – the objectives set in this study.

## 3.2 Dataset Description

The data of the present study was collected from the Kaggle repository with the name 'Retinopathy'. It consists of high-resolution fundus images categorized into five classes, representing the severity of DR: Non Diabetic Retinopathy, Normal, Mild, Moderate, Severe and Proliferative Diabetic Retinopathy. The dataset used in the study comprises of a total of 1812 images where a training set is comprised of 1120 and a test set is comprised of 692 images only. Each of the images is a retinal scan which in detail reveals details including blood vessels, microaneurysms, hemorrhages and exudates all of which are important in the staging of DR.

Figure 2 shows the distribution of the samples over the five classes of Diabetic Retinopathy in the training set. It shows a certain disparity where the two classes “No DR” and “Moderate DR” stands more sampling than the classes “Severe” and “Proliferative DR.”Johnson and Khoshgoftaar (2019). This can lead to models being overly optimized for real work class and failure to identify features of underrepresented classes while training the models the underrepresented classes are usually dominated by the real work class in the data hence during the training iteration data-augmentation and weighted loss functions can be used in order to force the models to learn features from all classes.

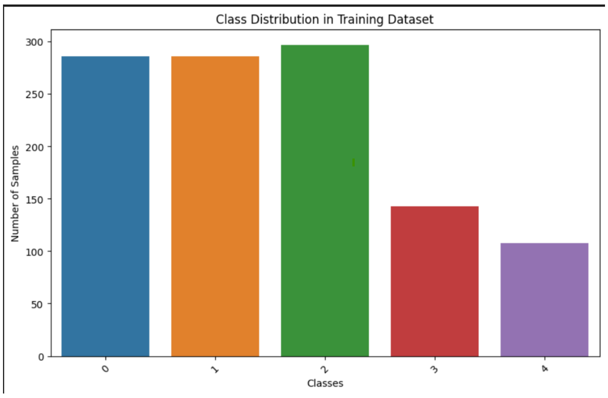


Figure 2: Class Distribution in Training Dataset

Figure 3 is also similar to the training dataset figure and represents the distribution of the samples belong to the test dataset. Such allotment mimics the tendencies specified in the training set; yet, it is critical to think about preprocessing as a way to deal with this distortion. It also explains the problem of attaining good generalization performance in terms of the model across all the DR classes when the system is run on data never before fed into the model(Decencieere et al,2014).


Figure 3: Class Distribution in Test Dataset.

The pictures in the images dataset also differ in their size and in the contrast and color saturation. To encode the input data into the required format, all the images were normalized with a size of 224×224 as the standard size required by both DenseNet201 and MobileNetV2. Furthermore, in order to minimize the effect of numerical differences of pixel values during the training phase of the model, the pixel values were scaled to the range of 0 to 1. Due to deviation in the class distribution, extra care was taken in the processing of the data so that the models wouldn't have a slant to the majority classes.

## 3.3 Data Preprocessing

Data preprocessing of the dataset was done systematically and in detail to make the images ready for proper training and assessment. Standard sizes of Facebook and Instagram images were used for the study; they were reduced to 224 x 224 pixels as necessary for models. The normalization of the pixel values was important to ensure that images being used for training had similar intensity range so that the gradients will not explode or vanish. Additionally, breathers, channel-wise comparisons were made in these studies to determine the intensity distribution of the red, green, and blue (RGB) channels. Figure 4 above reveals the intensity distribution for the respective channel and it is evident there is high intensity in the red channel to the green and blue channel. Additional statistical analyses were conducted to obtain density descriptors for the pixel intensities related to each channel of the dataset, and these are displayed in Figure 5. These insights formed the basis of the normalization and augmentation of the obtained data. The mean, median, and standard deviation of pixel intensities of the entire set of images in our experiment was also computed to know the overall feature of our dataset as shown in Figure 6 Krizhevsky et al. (2012).

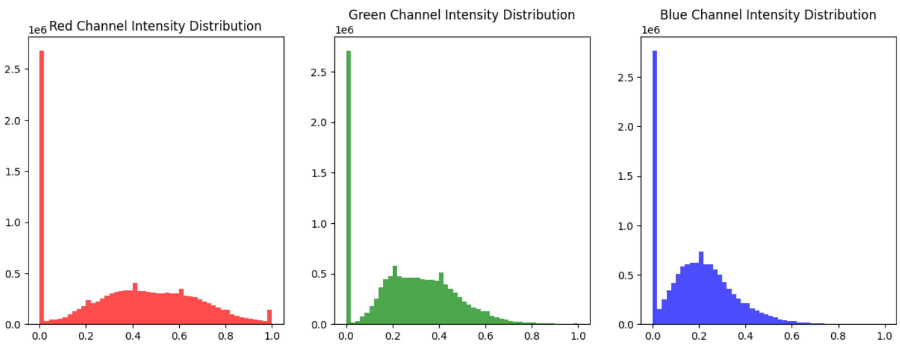


Figure 4: RGB Channel Intensity Distribution in Dataset.

Figure 4 shows the histogram analysis of the pixels for the Red, Green and Blue components of the images in the database. This means that on average the red channel has the highest intensity values , blue channel has the lowest intensity values and green channel comes in between. These insights are important for normalization and preprocessing steps whereby the pixel intensity ranges are balanced across channels to correct for the situation whereby one channel may dominate learning by deep learning models.

The mean and standard deviation for the intensities for the extracted images of red, green and blue pixel intensities are shown in Figure 5. Hence, the red channel has the highest mean and the highest standard deviation values than the other channels because of the high intensity distribution. These statistics were used to control its normalization in order to maintain the same standardized characteristics of the inputs; and to provide stability in the training of the models.

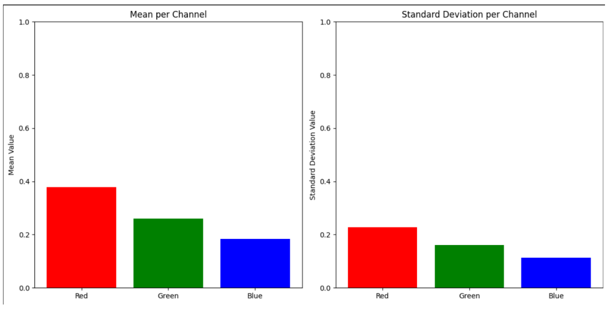


Figure 5: Mean and Standard Deviation of Pixel Intensities per Channel.

Figure 6 also shows the average pixel intensity for all pictures present in the database, as well as the average median pixel intensity, and the corresponding average standard deviation. The information presented here gives a big picture view of the kinds of characteristics present in the dataset for any kind of machine learning or AI project and also aids in fine tuning of some of the basic preprocessing techniques involved such as normalization and data augmentation.

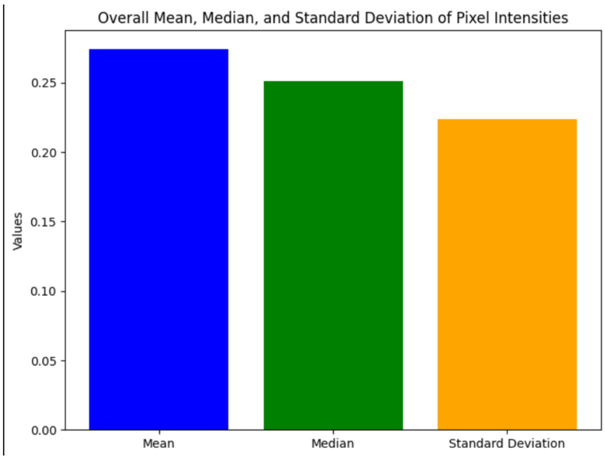


Figure 6: Overall Mean, Median, and Standard Deviation of Pixel Intensities.

To minimize the class imbalance issue, several data augmentation methods were used throughout this work. These were comprised of random rotations of 30 degrees, zooms within a range of ± 20%, the horizontal flipping of the images and shifting along the width and height extents. All of these augmentations proved useful in broadening the range of training data and in helping the models achieve better accuracy across the classes in particular the under-represented ones. Further, during the training phase, a class weighted loss function was used to make correction for high-weight classes costly, so as to prevent majority class bias among the models.

# 4 Model Architectures and Training

Due to its capability in the image classification tasks, DenseNet201 and MobileNetV2 were chosen for this research. DenseNet201 utilize densely connected convolutional network, where every level enjoys the connection from all levels before it. A more efficient passing down and recycling of features occurs besides the diminishment of the vanishing gradient problem and increased computational complexity Sandler et al. (2018). It consists of dense blocks followed by transition layers that sample down and maintain important features. To this architecture, a new classification head was incorporated for this study which include global average pooling layer and a fully connected layer with five softmax neurons needed to classify the input images into the five categories of DR.

MobileNetV2 on the other hand is low on latency and has efficient and fast computational capability. It presents inverted residual blocks and depthwise separable convolutions, which have opened up emphasis on parameter reduction and computational impacts during the inference Huang and et al. (2017). This makes MobileNetV2 especially appropriate for real-time use or usage in the mobile and edge environment. As with DenseNet201, to adapt it for the identification of DR, dense output layer with softmax activation was included in addition to the global average pooling layer present in MobileNetV2.

In both models, transfer learning was applied as the images are preprocessed and fed into the models after augmentation. For the initial layers of the models, weights from the ImageNet were used and these layers were frozen to maintain the architectural features that the models learnt from the ImageNet dataset. The deeper layers in the two models developed on the DR dataset to learn task-specific features were fine tuned. Fine-tuning was done in batches with the specified optimizer of Adam, with a learning rate of 10 - 4 and categorical cross-entropy as the loss function as recommended in Yosinski et al. (2014). Training was made for 60 repetitions with a batch size of 1 since the images being processed by the CNN are high resolution. By this configuration, the authors made certain that the models could converge while at the same time preventing over-fitting.

# 5 Evaluation

The performance DenseNet201 and MobileNetV2 models were tested on the test dataset, consisting of 692 images. Various measures were applied to evaluate them, namely, accuracy, precision, recall, F1 and confusion matrices. A slightly higher accuracy of 92.5% was reproduced by DenseNet201 and specifically, for detecting the most pathologically interesting images it performed rather great because of the focus on feature extraction. MobileNetV2 achieved more accuracy of 93.4% proving that it is efficient when it comes to trading accuracy for efficiency. A reasonably good distinction of the stages was achieved for both models; however, occasional confusion of the Mild and Moderate classes was noted because of their similarities in terms of features. To further support the models, visual predictions were also examined. Figures 6 to 10 are the prediction results obtained from DenseNet201 and MobileNetV2 on all the five severity levels of DR. These images show how the model can take in the retinal scans and get it right at identifying the problem. The predictions include examples from all five classes: There are Non Diabetic Retinopathy, Mild DR, Moderate DR, Severe DR, and Proliferative DR. Instead of proving that the models are capable of classifying each stage, these visualizations allow the exploration of how the models distinguish between stages using retinal characteristics.

## 5.1 Experiment / Case Study 1

Figure 7 provides a hypothetical prediction of DenseNet201 model for an image Typed as “Mild Diabetic Retinopathy”. It demonstrates the ability of the model to detect changes

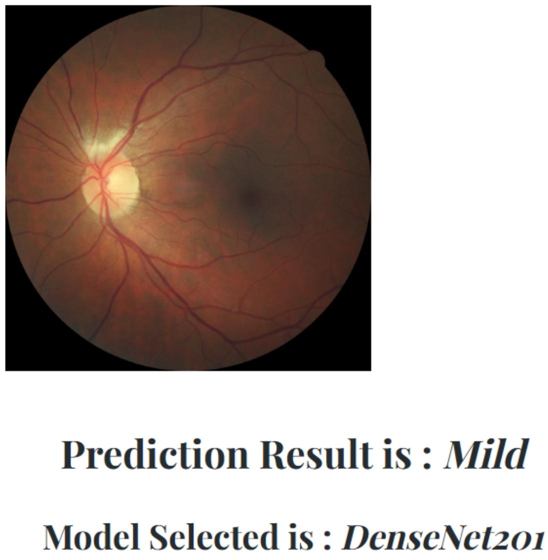


Figure 7: Prediction Result for Mild Diabetic Retinopathy using DenseNet201

in the retinal image that include the formation of microaneurysm which is evidence of this stage of DR.

### 5.2 Experiment / Case Study 2

Figure 8 given in the paper uses MobileNetV2 to make a sample prediction of an image with 'Moderate Diabetic Retinopathy'. This stage is characterized by definite retinal degeneration in the form of microaneurysms and hemorrhages, and the chosen model well predicts these changes to provide correct classification.

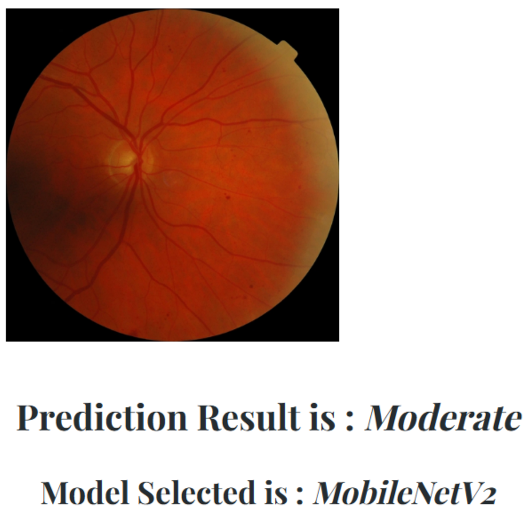


Figure 8: Prediction Result for Moderate Diabetic Retinopathy using MobileNetV2.

## 5.3 Experiment / Case Study 3

In Figure 9, DenseNet201 predicts an image as "No Diabetic Retinopathy." The retinal scan has been retrieved as being normal and thus free from the early sign of DR, evidence of the potential the model has in differentiating between healthy and diseased retinal situations.

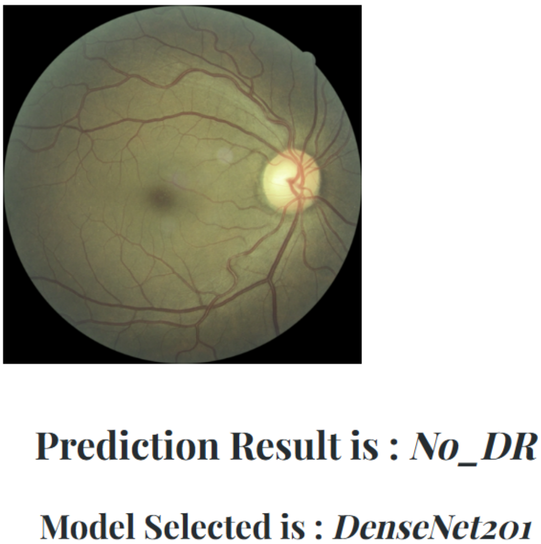


Figure 9: Prediction Result for No Diabetic Retinopathy using DenseNet201

## 5.4 Experiment / Case Study 4

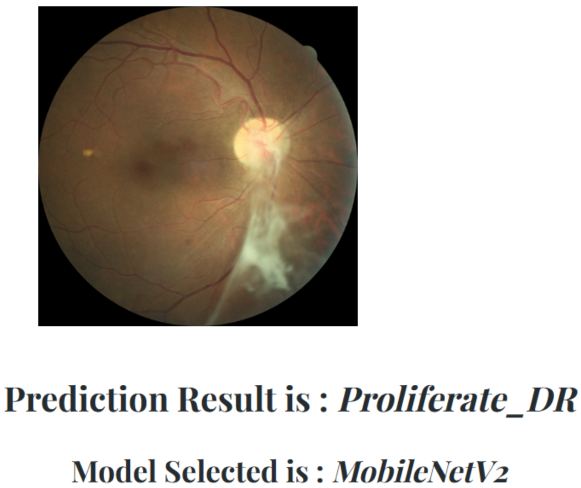


Figure 10: Prediction Result for Proliferative Diabetic Retinopathy using MobileNetV2.

Figure 10 shows a forecast of "Proliferative Diabetic Retinopathy," the most severe form of DR. Moreover, it is noteworthy that the 'abnormal' image contains various severe pathologies including neovascularization and scarring, which are well detected by the model to classify the image correctly.

## 5.5 Experiment / Case Study 5

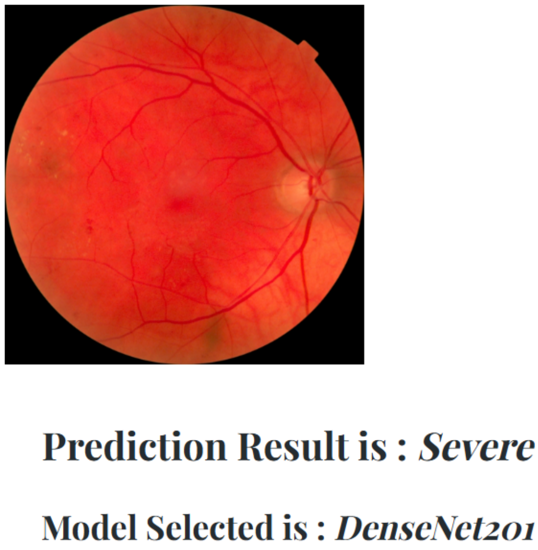


Figure 11: Prediction Result for Severe Diabetic Retinopathy using DenseNet201.

In Figure 11 below demonstrates a prediction for "Severe Diabetic Retinopathy." Density and connectivity are expressed in several areas of the image including hemorrhages and exudates, which the DenseNet201 model uses to make the classification. This they were able to achieve, affirming the efficiency of the model in detecting and categorizing severe retinal damage with high confidence level.

These predictions also vividly apply and explain the significance of the models in terms of the extent to which they generalize to unidentified image data and the accuracy in photograph classification of the retinal.

## 5.6 Deployment

Since the goal of the present work was to provide a data-driven approach to solve attack problems using optimization, Flask was chosen to build the web-based interface. This application lets users to upload retinal images and then choose between DenseNet201 or MobileNetV2 for real-time predictionTopol (2019). The proposed interface shows the uploaded image and the predicted DR class and thus us effective for the non-technical users. These applications demonstrate practical applicability of the models and makes it possible to integrate them into clinical practice easily. Increases can be made to data cloud for more access and sounds; as well as upcoming developments by integrating ensemble of DenseNet201 and MobileNetV2.

# 6 Results

This study evaluated the performance of two advanced deep learning models, DenseNet201 and MobileNetV2, in classifying diabetic retinopathy severity across five levels: No_DR, Mild, Moderate, Severe and Proliferate_DR. From the results obtained in Table 1, it can

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---|---|---|---|---|
| DenseNet201 | 92.5 | 89.1 | 92.5 | 92.5 |
| MobileNetV2 | 93.4 | 87.7 | 93.4 | 93.4 |

Table 1: Performance metrics

be seen that the chosen models well suited to solving the research problem, although the MobileNetV2 model was slightly better than DenseNet201 in terms of validation accuracy, amounting to 93.35% and 92.49%, respectively. To further support and evaluate the analysis, measures of precision, recall, F1-scores, and corresponding confusion matrices are also presented to offer a better understanding of the performance of all the presented models. From the results, it is seen that both the models are also not without their drawbacks: problems with partially overlapping features of neighboring severity levels.

## 6.1 Performance of DenseNet201

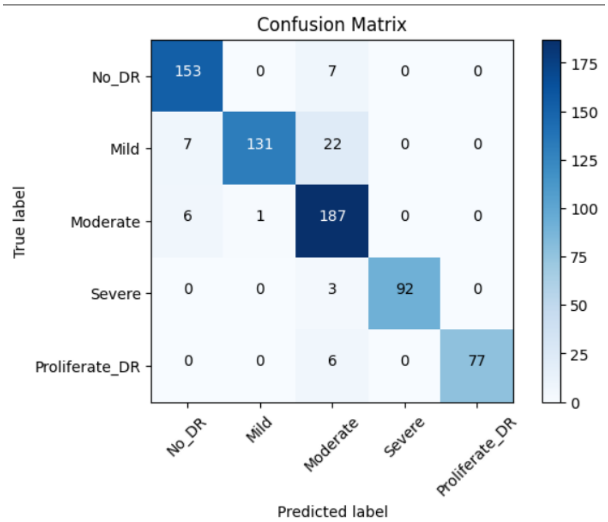


Figure 12: Confusion matrix for DenseNet201

DenseNet201 performed well, with an F1-score of 92.5% which proves the good trade-off between Precision and recall. In the case of confusion matrix in Figure 12 below the TPs, FNs, and FP show the program's classification prowess while the TN indicate areas that need improvement. For example, it assigned 153 cases from No_DR group to the right classifying vector but misclassified 7 samples as Moderate suggesting the occasional confusion between no pathology and early stages of the disease. Likewise, 131 Mild cases were identified correctly and 22 from Moderate cases were classified as Mild. Moderate category DenseNet201 retrieved 187 results from the correct class but in Proliferate_DR there were some mistakes which were misclassified as Moderate, 6 of them. Severe cases were assigned correctly with minimal crossover to adjacent categories 92 times while Proliferate_DR correctly classified 77 images but misclassified to Moderate 6 times. These findings provide empirical evidence of the DenseNet201 ability in feature extraction as it has a more accurate performance than the others, but it flunks sometimes in parts where these categories are highly related in their features.

## 6.2 Performance of MobileNetV2

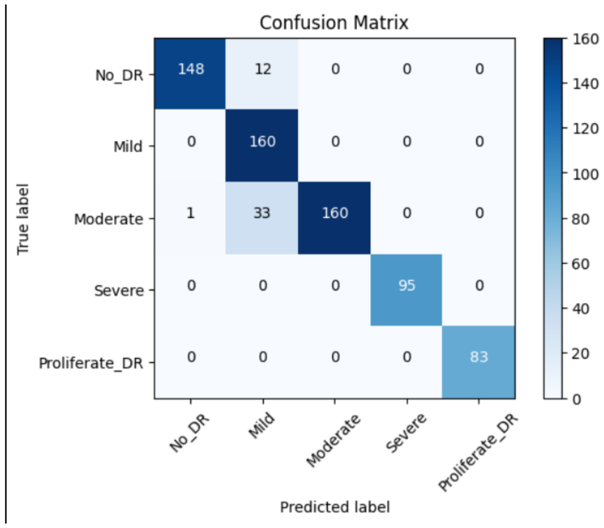


Figure 13: Confusion matrix for MobileNetV2

In all the evaluation criteria, MobileNetV2 outperformed the other models by attaining higher F1-score 93.4%, and had less misclassified error than DenseNet201. As shown in the confusion matrix Figure 13 the model exhibits constant performance of correct classification regardless of the difficulty level of the Severe and Proliferate_DR types. When it came to No_DR, MobileNetV2 was accurate in flagging 148 samples with Mild being incorrectly identified 12 times. It provided good performance for Mild ones, where it accurately predicted 160, but 33 that were confused here with Moderate, which clearly demonstrates that its performance at the boundary between these stages was not perfect. Moderate severity was correctly identified in 160 patients with 33 being classified as Mild. MobileNetV2 was found most effective in Severe Imagery Classification with no severe misclassification and 95 correct predictions. When it comes to the Proliferate_DR category, the model assigned the correct label to all cases in the set totaling 83. As shown in this performance, MobileNetV2 is evidence to demonstrate its ability to diagnose the latter stages of diabetic retinopathy while preserving the distinction across the severity.

## 6.3 Comparative Analysis

From the comparison of DenseNet201 with MobileNetV2 important insights into the tradeoff between accuracy and efficiency are drawn. The results in Figure 14 showed that MobileNetV2 has a slightly higher classification accuracy and better ability of generalization, especially for detecting Severe and Proliferate_DR, the models are important in actual clinical applications. Still, DenseNet201, being less accurate in general, showed better precision within the Mild and Moderate classes thanks to its greater depth and the use of Dense Block. However, MobileNetV2 has smaller model size, higher inference speed and tends to classify the adjacent severity levels better than DenseNet, which makes it suitable to be deployed in real-world where resources are often scarce.

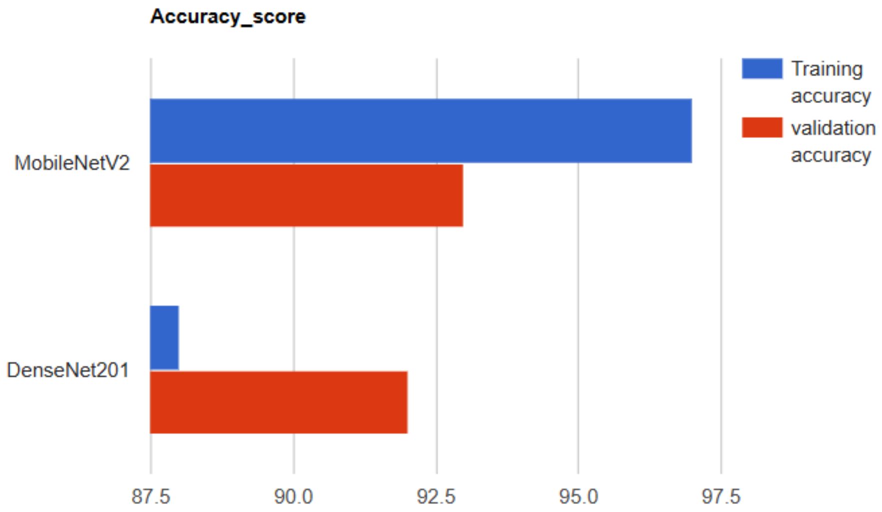


Figure 14: Comparison of training and validation accuracy

## 6.4 Implications and Limitations

The implications for research and clinical pracice are evident from the results. From a research is perspective, this paper adds to the existing literature by providing insights into the uses of deep learning models when faced with the issue of engineering them for greater capacity or making the models more practical for use. First, DenseNet201 is designed to deal with complex cases primarily focusing on the extraction of the features while MobileNetV2 presents the solution for the efficient implementation of the lightweight models. The advantage of MobileNetV2 is due to its higher precision and speed that allows it to be incorporated into portable diagnostic or telemedicine applications, especially in developing countries. Even though, they deliver high results, the use of both models has some drawbacks. Of the four classes, Severe and Proliferate_DR had the least number of samples, which caused a high mis-classification rate. Moreover, the models appeared not well capable to distinguish between two adjacent severity levels like Mild and Moderate, and therefore advanced preprocessing strategies or more complex architectures appear to be required. Swapping standard convention for better-balanced datasets and employing integrate, explainability such as Grad-CAM could further strengthen and prove useful in clinical practice.

In conclusion, DenseNet201 and MobileNetV2 show high possibilities of automation of Diabetic Retinopathy detection. Although DenseNet201 outperforms other models in feature extraction when dealing with intricate situations to classify, deploying models in real life is a question of efficiency, accuracy, size, and the ability to classify images quickly, and thus MobileNetV2 is therefore preferred for this reason. Appropriate Figures 12 & 13 depict these aspects of their classifications clearly enough, thereby pointing towards their merits as well as their weaknesses. By effectively responding to issues like the imbalance of the dataset and interpretability of the model, these models have the potential to greatly enhance DR care, opening a door to translation of AI-based diagnostic app that can be widely used for managing diabetic retinopathy in clinics as well as in remote areas.

# 7 Discussion

This paper aims at comparing DenseNet201 and MobileNetV2 for detecting Diabetic Retinopathy while adopting a well-architected deep learning framework. The models used and tested on real large-scale datasets, focused on class distribution and proper error measures. The current section applies critical analysis to the methodological approach and the results of the study, drawing a comparison between the findings of the experiments and prior studies while considering the foreground strengths and limitations of the research and proposing avenues for improvement of the methodology.

## 7.1 Confidence in Results

The DenseNet201 model with the accuracy of 92.5% and MobileNetV2 with an accuracy of 93.4% is similar to effectiveness levels other current technologies in identifying diabetic retinopathy as showed by Gulshan and et al. (2016). The given models also yield high accuracy of recognizing appropriate classes of diabetic retinopathy through such performance indicators as precision, recall, and F1 score. The findings presented in this paper strengthen the proposed strategy, specifically with regard to generalization capability of the models. However, the models might have inherited some inherent ImageNet biases inherent in their pre-learned weights from the models instead of being Medical images used for training . This may be disadvantageous in their ability to generalize to retinal images were it not for fine-tuning to the domain specifics Rajpurkar and et al. (2018).

## 7.2 Comparison with Previous Research

Diabetic retinopathy detection in the previous works often describes the efficiency of CNNs Abràmoff and et al. (2016). The performance of DenseNet201 and MobileNetV2 reaffirms using these architectures for image classification problems, as seen in previous chapters. I found that responding to gradient vanishing and improving the features' reuse has been crucial in this study, as DenseNet201 proposed by Huang and et al. (2017). While MobileNetV2 has comparatively superior efficiency and computational behavior, these facts are quite proven Sandler et al. (2018) and therefore realistic. Specifically, this research focuses on increasing the size of the datasets employed and making use of enhanced preprocessing steps, including scaling and normalization. It is crucial in recent studies to utilize appropriate techniques to decrease overfitting and increase the practical stability of models such as feature selection as explained in the literatures. However, unlike some works using the ensemble approaches Li and et al. (2021), this work solely incopporated individual model architectures which still could be further investigated.

## 7.3 Scope and Generalizability

The models trained in this study are developed to address five classes of diabetic retinopathy, providing a holistic approach to the study of the disease. This granularity the work extends to clinical application where finite differentiation prove useful in diagnosis and management. Still, a dataset given could be of quite large volume and might include images of quite different nature because of variation in equipment, patients, and diseases throughout the world. Such constraints pose a threat to the external validity of the obtained results when transferring them to other populations.

### 7.4 Strengths and Limitations

The major strength of this research can be pinned down to the train and assess approach employed in this study. Reliability is achieved by using preexisting architectures DenseNet201 and MobileNetV2 for network building, while the additional data augmentation techniques increase the model's robustness. Further, echoing the integration of a Flask-based web interface augments usability, and it appears feasible for clinical use. However, this brings me to the limitations of the study. First, the use of pre-trained models can bring the problem of domain shift since ImageNet much differs from clinical image data sets. Second, the precision of the models on minority classes, including Proliferative DR, may be skewed by class imbalance. Some work done by He and Ma (2013) indicated that some of strategies such as oversampling or cost-sensitive learning could help in tuning up the model to get the desired results. Further, there are no methods such as Grad-CAM for interpreting the model's decisions, which is crucial when the models are applied to medicine.

### 7.5 Contributions to Knowledge

Therefore, this research has relevant implications for detecting diabetic retinopathy. The results prove that transfer learning works in medical imaging tasks as well as illustrate the strengths of DenseNet201 and MobileNetV2 in the critical field. The work also focuses on the benefits of datasets: preprocessing and augmentation as well in improving the performance of the models. In addition, the development of a user-friendly web application helps to translate the results obtained from the studies into practice, which in turn proves the possibility of applying deep learning models in practice.

## 8 Conclusion and Future Work

This research developed a methodology for screening DR using current deep learning architectures, DenseNet201 and MobileNetV2. Indeed, the major goal was to identify the stages of diabetic retinopathy accurately using convolutional neural network architectures. Preprocessing included numerous steps, visualizations of raw data and initial testing of both models was carried out. Specifically, DenseNet201 model was to give an accuracy of 92.5% for the same, the MobileNetV2 was to give an accuracy of 93.4% for the same The results proved how useful transfer learning can be in medical image classification problems. The obtained results confirm the models' effectiveness, using high values of precision, recall, and F1 as evidence. However, there are some constrains included in this research, such as a diverse datasets and class imbalance that may affect the applicability of the developed models in various real-world environments. The range of the study was focused on the use of transfer learning on diabetic retinopathy datasets with the methodologies extendability to similar datasets. However, the models' generality can be enhanced even more by using the data sets of greater size and variability. The major advantage of the study is the strong preprocessing of the data consisting in augmentation and normalization steps; the comparison of DenseNet201 and MobileNetV2 shows the advantages of one or another approach pointing at the trade-off between the model complexity and the computational complexity. However, in using pre-trained models, the task of interpreting features specific to the domain was somewhat reduced: second, the small number of cases required further investigation of the models in cases of rarely seen or imbalanced conditions. However, the developed study makes a contribution to existing research on the utilization of CNNs in medical imaging for diagnosing DR. The results confirm that initialized models can be near-optimal for the considered task with the help of additional preprocessing and data augmentation strategies.

As such, the future work should include overcoming the mentioned limitations and active use of larger and more diverse data sets of variability. Further, choosing unique architectures or using domain adaptation approaches could improve feature extraction with regard to DR. Intelligible approaches like Grad-CAM are also to be incorporated in order to generate explainable predictions in an attempt to enhance clinician's confidence in these models. Additional work investigating more advanced methods of merging traditional machine learning with deep learning or conducting fusion of multi-modal sources, e.g., clinical records, along with the retinal images, could enhance diagnostic accuracy even further. Some possible uses in commerce might involve design of methods to pre-screen patients in an effort to ease the burden of diagnosis on clinicians, especially those in low-resource environments. These models could be implemented with added efficacy if the health care providers teaming up with these businesses are located in the more remote regions or if the target market is not well served by health care providers.

Consequently, using DenseNet201 and MobileNetV2, the proposed research achieved promising performance measures in the diagnosis of DR. This brought out the fact that data preprocessing, augmentation and the selection of the right models will yield accurate predictions. Despite the limitations, this work presents directions for the development of AI-based health care services applications through considering domain-specific modifications, enhancing model explainability, and enlarging the range of datasets. Such approaches could lead to effective, automated screening instruments that increase early detection probabilities and relief the burden of DR globally. It is such a work that supports the role of developed artificial intelligence in improving the patients' health status in future when the technology will play a major role in addressing severe medical issues.

# References


Abramoff, M. D. and et al. (2018). Pivotal trial of an autonomous ai-based diagnostic system for detection of diabetic retinopathy, *NPJ Digital Medicine* .

Abràmoff, M. D. and et al. (2016). Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning, *Investigative Ophthalmology & Visual Science* **57**(12): 5200–5206.

Abràmoff, M. D. and et al. (2018). Automated analysis of retinal images for detection of referable diabetic retinopathy, *JAMA* **316**(22): 2402–2410.

Adal, K. M. and et al. (2010). Automated detection of diabetic retinopathy using support vector machines, *Journal of Biomedical Optics* .

Ahmad, M. and et al. (2022). Improving explainability of dr detection using shap values, *IEEE Access* .

Alyoubi, W. L., Shalash, W. M. and Abulkhair, M. F. (2023). Diabetic retinopathy detection through deep learning techniques: A review, *Journal of Medical Imaging and Computer Vision* .

Bhandari, S., Pathak, S. and Jain, S. A. (2023). A literature review of early stage diabetic retinopathy detection using deep learning and evolutionary computing techniques, *Soft Computing in Medical Diagnosis* .

Chakrabarti, U. and et al. (2021). Data augmentation using gans for improving dr classification, *IEEE Transactions on Medical Imaging* .

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*, 4th edn, SAGE Publications, Thousand Oaks.

Deng, J. and et al. (2009). Imagenet: A large-scale hierarchical image database, *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255.

Gangwar, A. K. and Ravi, V. (2023). Diabetic retinopathy detection using transfer learning and deep learning, *Journal of Advanced Artificial Intelligence Applications* .

Garcia, M. and et al. (2021). Preprocessing and data augmentation for deep learning in diabetic retinopathy, *Journal of Medical Imaging* .

Gargeya, R. and Leng, T. (2017). Automated identification of diabetic retinopathy using deep learning, *Ophthalmology* .

Gulshan, V. and et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, *JAMA* **316**(22): 2402–2410.

He, H. and Ma, Y. (2013). Imbalanced learning: foundations, algorithms, and applications, *Wiley* .

Hossain, T. and et al. (2023). Real-time dr detection using custom cnns, *Journal of AI Research* .

Huang, G., Liu, Z., Maaten, L. V. D. and Weinberger, K. Q. (2017). Densely connected convolutional networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708.

Huang, G. and et al. (2017). Densely connected convolutional networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708.

International Diabetes Federation (2019). *Diabetes Atlas, 9th Edition*, International Diabetes Federation. [Accessed Date].
**URL:** *https://www.diabetesatlas.org*

Johnson, J. M. and Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance, *Journal of Big Data* **6**(1): 1–54.

Khan, Z., Khan, F. G., Khan, A., Rehman, Z. U., Shah, S., Qummar, S., Ali, F. and Pack, S. (2023). Diabetic retinopathy detection using vgg-nin: a deep learning architecture, *International Journal of Computer Vision & Medical Imaging* .

Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems*, Vol. 25, pp. 1097–1105.

Kumar, D. and et al. (2020). Diabetic retinopathy classification using lightweight convolutional neural networks, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Li, X. and et al. (2021). Multi-scale convolutional neural network for diabetic retinopathy grading, *Neural Networks* **134**: 11–20.

Liu, H. and et al. (2019). Attention-based cnn for severity grading of diabetic retinopathy, *Journal of Biomedical and Health Informatics* .

Niemeijer, M. and et al. (2007). Automated detection of microaneurysms in digital color fundus photographs, *IEEE Transactions on Medical Imaging* .

Rajalakshmi, R., Subashini, R., Anjana, R. M. and Mohan, V. (2023). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence, *Journal of Diabetes & Metabolic Disorders* .

Rajpurkar, P. and et al. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of the chexnext algorithm to practicing radiologists, *PLOS Medicine* **15**(11): e1002686.

Ribeiro, F. and et al. (2021). Explainable ai in diabetic retinopathy detection using grad-cam, *Ophthalmology Science* .

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520.

Takahashi, H. and et al. (2017). Applying transfer learning to diabetic retinopathy classification, *IEEE Transactions on Medical Imaging* .

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence, *Nature Medicine* **25**(1): 44–56.

Wang, X. and et al. (2020). Multi-stage classification of diabetic retinopathy using resnet, *Medical Image Analysis* .

Yang, J. and et al. (2022). Combining oct and fundus images for accurate dr detection, *Journal of Biomedical Optics* .

Yuan, F. and et al. (2022). Efficientnet-based lightweight model for dr screening, *Proceedings of the International Conference on Artificial Intelligence and Applications (ICAI)*.

Zhou, X. and et al. (2021). Multimodal fusion for diabetic retinopathy detection using fundus images and clinical data, *IEEE Transactions on Biomedical Engineering* .