Diabetic Retinopathy Stage Detection using Ensemble EfficientNetV2B3 and NasNetMobile

1[st] Md. Anas Mahmud
ID: 20101149

2[nd] Navid Hasan Rafi
ID: 20101585

3[rd] Tajrian Mahbub
ID: 20101325

***Abstract*—Our research project focuses on using machine learning and medical data to create an accurate model for predicting diabetic retinopathy. By analyzing patient's eye images, we employ convolutional neural networks (CNNs) to classify the disease stages. Our goal is to achieve a low-cost method for early detection by selecting the best-performing deep learning architecture. We collected 3663 train-samples and 1929 test-samples from Kaggle's APTOS'2019 Blindness Detection Competition. The model should be capable of classifying diabetic retinopathy stages from mild to severe based on the eye images. We plan to utilize Sequential, ResNet152V2, InceptionV3, Xception, DenseNet201, EfficientNetV2B3, NasNetMobile and a hybrid of EfficientNetV2B3 and NasNetMobile along with image preprocessing techniques, to improve the model's accuracy. We also address potential challenges, such as limited dataset availability, image quality issues, and dataset imbalance, proposing solutions like data sharing and image preprocessing to overcome these challenges. Our work aims to contribute to the advancement of diabetic retinopathy detection systems using deep learning techniques.**

I. Introduction

Diabetic retinopathy, an incapacitating complication arising from diabetes mellitus, emerges as a prominent contributor to visual impairment and blindness on a global scale. The adverse effects of this progressive retinal disease, if left unattended, can significantly compromise the overall well-being and daily functioning of the individuals it afflicts. The importance of timely diagnosis and intervention cannot be overstated, as it plays a crucial role in mitigating the devastating consequences associated with the subject under consideration. In the modern age, characterised by the emergence of artificial intelligence (AI) and deep learning, a paradigm shift is observed in the healthcare sector. This shift entails a transformative journey aimed at harnessing the immense potential of state-of-the-art technology in the field of healthcare. The focus of our research project revolves around the advancement of a groundbreaking solution aimed at detecting diabetic retinopathy. This solution has the capacity to significantly transform the process of screening and diagnosing this condition. The present endeavour embodies the amalgamation of medical proficiency and machine learning acumen. The objective of this study is to develop a reliable and precise diagnostic tool for the swift identification of diabetic retinopathy from retinal images, utilising cutting-edge deep learning models and advanced computer vision techniques. The primary objective of our organisation is to improve the efficiency and accuracy of diagnostic procedures, with the ultimate goal of promoting timely intervention and preserving visual health. In the subsequent sections, an in-depth analysis is conducted on the intricacies of the approach, wherein the methodologies, models, and data supporting this ambitious endeavour are thoroughly examined. In the pursuit of enhancing the detection of diabetic retinopathy, a condition that poses a significant threat to vision, our research endeavours involve rigorous experimentation and validation. Our primary objective is to shed light on a more efficient, accessible, and effective approach to identifying diabetic retinopathy. By doing so, we aim to enhance the quality of life for individuals who are susceptible to this potentially debilitating ailment. This paper invites readers to embark on an exploration of the convergence between the fields of medicine and machine learning. Specifically, it delves into the intricate realm of diabetic retinopathy detection, aiming to shed light on the challenges and intricacies associated with this condition. The ultimate goal is to envision a future where early diagnosis of diabetic retinopathy is universally accessible, thereby preserving invaluable eyesight and revolutionising the healthcare landscape.

II. Background

We have collected about 3663 train-samples and 1929 test-samples from Kaggle's APTOS'2019 Blindness Detection Competition. The training data is used to train the deep learning model to learn patterns and features associated with different stages of diabetic retinopathy. The test data is used to evaluate the model's performance on unseen samples and assess its ability to generalize to new data. Our goal is to use deep learning to detect diabetic retinopathy and classify its stages. The model should be capable of analyzing the eye images and classifying them into different stages of diabetic retinopathy, ranging from mild to severe.

III. Ideas

Our main idea consists of what our research project is going to achieve. For easier to understand, our main idea consists of two parts. Now before going into the sub parts, we need to discuss how our program will work. First of all, diabetic retinopathy is a disease where the patient suffers from eye disease because of diabetics itself. When the patient has this disease their blood vessel leaks and forms blood around them. So what we want is to capture the images of the patient's eyes and put them into our CNN model.

A. Sub Task 1

The first thing we want our deep learning program is to detect the diabetic retinopathy itself. The main classification contains of two types of data, the eyes with and eyes without diabetic retinopathy disease. The Models will look through each images and see if they have formed blood around their blood vessel in a circular way. The model will also look if they have an increased amount of blood vessel around their retina which also indicates positive for the disease.

B. Sub Task 2

Since diabetic retinopathy has 4 stages, which are Mild, Moderate, Severe and Proliferative diabetic retinopathy. Mild is the earliest stage which has a few tiny swelling around the blood vessel and micro fluids can be seen at the center of the retina. Moderate has increased amount of tiny swelling and Severe stage has larger section of blocked blood vessel in the retina. The last stage Proliferative has few advanced features such as thin but many blood vessel, large section of blood forming around the abnormal blood vessel. We want our research project to be able to classify the stages of diabetic retinopathy by extracting those features.

IV. Plans

Now this section contains the things we plan to do with our project. It contains basic sub parts for better understanding.

A. Data Collection

As mentioned above we got our data set from a reliable competition which is from Kaggle Competition named "Aptos Blindness Detection 2019". There are about 5,590 images with which is about 10GB but we will use about 3662 images for our train case here as they are labled. We did want to use the test set as well but they were not labled and only meant predict the labels after training. All the images are labled on 5 types of integer ranging from 0 to 4. 0 means no diabetic retinopathy and 1 to 4 contains the stages of diabetic retinopathy. The 3662 data points offered for training are divided in two columns - unique value (file name of .png files) and diagnosis. Here, our most significant column is the unique values as we need the pictures for the whole process of detection and classification. Then, 'diagnosis' column labels the pictures according to the stages of diabetic retinopathy. The dataset's goal is to detect diabetic retinopathy and classify the stages based on pictures. There is no mismatched and missing data in the train set. The mean and std. deviation of the train data are 1.13 and 1.3. Maximum pictures of the train set is labeled as stage 0.

B. Model Training

We plan to choose a lot of model and even the latest ones such as ConvNeXt which was released in 2020 and also a lot of Keras Pretrained models ranging from popular ones to least used ones to test the accuracy. We also have a plan to use sequential model which uses CNN and optimize it accordingly to match the test accuracy of other popular models.

C. Improving Outcome

Our last but most important main plan is to improve our accuracy of our model. To do that we will be doing practical works such as rotating, flipping, scaling the images to a certain degree. Cropping the images to take the black border out to decrease noise. We also plan to use resized images which will take time for training but it can give out better results. In worst case we will change our models to improve accuracy while keeping the hybrid of two models unique.

V. Methodology

A. Data Acquisition And Preprocessing

The dataset was obtained from the Kaggle APTOS 2019 Blindness Detection competition and contains a variety of retinal pictures showing different stages of diabetic retinopathy. Firstly the data had many problems such as, some images were dark, some images were not correctly cropped and some images had different type of colors compared to other images. So we have completed our image processing using 2 steps and has been divided into 2 stages shown on the figure below with all 4 stages of diabetic retinopathy diseased eye.

1) The stage here shows the original images, which are not cropped, not properly sized to the border and some of the details such as veins and clogged wools can not be seen clearly.
2) We have changed the color of the images so none of the images would be dark and have same amount of light illumination on them.
3) On this stage, we improved the lighting of our images with Gaussian Filter and flattened the images and we also resized the images with our auto cropping function using color filters.

Dataset Segmentation: - Created three distinct subsets from the dataset: the training, validation, and test sets. The ratio of training, validation, and testing followed the standard distribution of 80% training, 10% validation, and 10% for testing.

Training the Model: - Initialized Sequential, ResNet152V2, InceptionV3, Xception, DenseNet201, EfficientNetV2B3, NasNetMobile and a hybrid of EfficientNetV2B3 and NasNetMobile with weights obtained from a broader context of image classification that have been pre-trained.

- Added a custom classification layer on top of the base model to correspond with the specific number of classes corresponding to the various stages of Diabetic Retinopathy.

- For training objectives, a suitable loss function, such as categorical cross-entropy, and Adam as the optimizer were chosen.

- To prevent overfitting, gradually train the model on the training data while continuously monitoring validation performance. We also included L2 Regulizers and Dropout layers to prevent overfitting.

Implemented an early prevention mechanism, terminating training when validation performance reached an end point .

Performance Assesssment: - We calculated the performance of each models based on how each models performs on the test set. We also included the train accuracy to show the comparisons.

Results and Analysis: - Result will include 6 pre-trained models by importing keras and it will also include the results of sequential model and a hybrid showing a comparison between them.

Discussion and Conclusion: - Engaged in an exhaustive discussion regarding the achieved results, the robust characteristics of the model, and potential avenues for improvement.

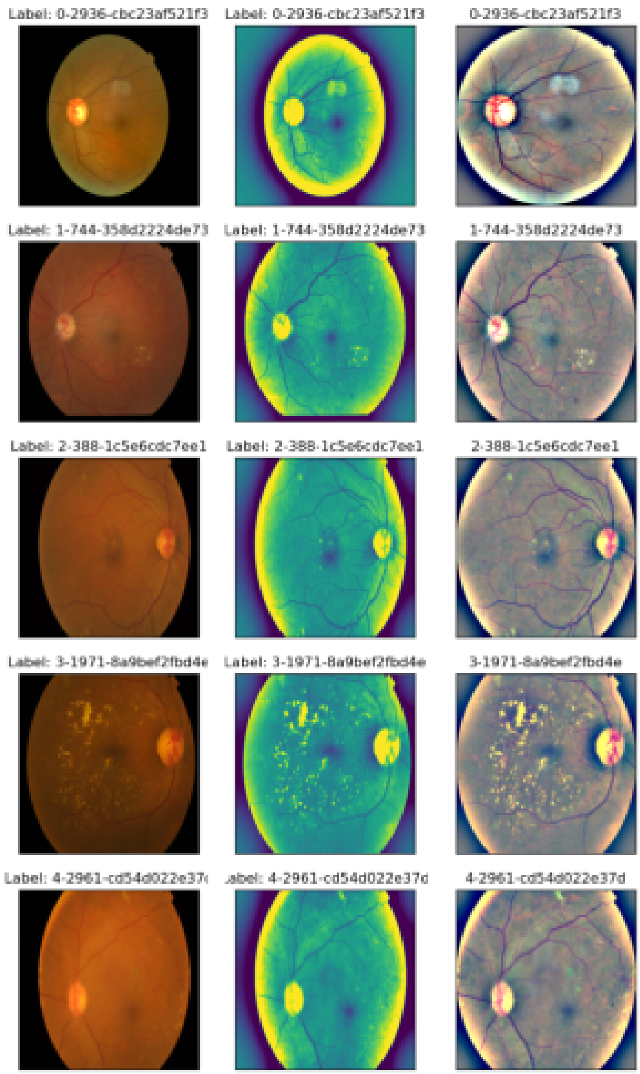

Fig. 1. Stage from 0-2

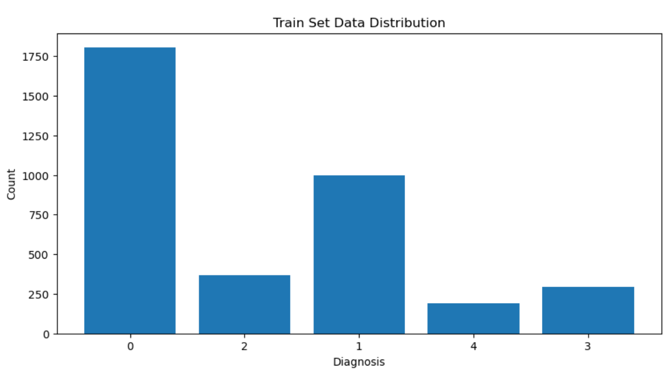

Fig. 2. Image Distribution between all stages

VI. Literature Reviews

Looking at a wide range of other papers regarding diabetic retinopathy, their approach and conclusion are rather interesting. In 2021, S. Shekar, N. Satpute, and A. Gupta wrote the initial paper. They made a detailed analysis of other papers, and among the 4 datasets they have used, DR positive and DR negative accuracy were 93.3% and 95.5%, respectively. In the second paper by L. Dai and others, they used Res Net to group the models, which they call DeepDR systems. These systems have three neural subnetworks. They had an F score of 90.1% and an AUC accuracy of 93.4%. The third paper consists of using RA-EfficientNet7, which comes up with an accuracy of 98.36% by San-Li Y.. The fourth paper by Z Khan consists of using the VGG-NIN model, which has an accuracy of 85% and an F score of 59.6%. We also had plans to evaluate another batch of papers, and among them was one by Sayantan, which was similar to the first paper mentioned above. The paper had an accuracy of 95.93 by using their proposed method, named DCNN. Furthermore, another outlined paper by Suwarna Gothane claims they used ResNet and came up with 82% accuracy. The last paper we would like to mention, is very latest by J. Pradeep, achieved an accuracy of 97.7% using CNN. They also mentioned testing SVM, which only had an accuracy of 48% in their test.

VII. Proposed Models

1. Xception: The Xception model, which was proposed by François Chollet in the 2017 research paper titled "Xception: Deep Learning with Depthwise Separable Convolutions," can be seen as a progressive advancement of the Inception architecture. The focus of this study lies in the utilization of depthwise separable convolutions, which aims to improve both computational efficiency and model performance. Xception is founded on the idea of depth-separable convolutions. The traditional convolutional layer is divided into two distinct phases by these convolutions: a depthwise convolution that processes individual input channels and a pointwise convolution that combines their outputs. Depth-wise separable Convolution: - Depthwise Convolution: Each input channel is filtered separately to capture spatial information. - Pointwise Convolution: Combines the outputs of the depthwise convolution using 1x1 convolutions, thereby capturing cross-channel information. Due to its ability to effectively capture spatial and channel-wise data, Xception retains a high level of representational capacity.

2. ResNet152V2: In 2016, He et al. created the convolutional neural network (CNN) ResNet152V2. It is a 152-layer deep CNN that performs exceptionally well on image classification tasks. ResNet152V2 is effective because of the hierarchical structure it uses to extract features from photos. Low-level characteristics, such edges and forms, are extracted in the network's initial layers. Higher-level characteristics, such as objects and patterns, are extracted by the network's later layers. ResNet152V2 can be used to extract features from fundus images for the purpose of detecting diabetic retinopathy. Images of the retina, the light-sensitive tissue at the back of the eye, are called fundus photographs. Fundus scans can reveal abnormalities in the retina caused by diabetic retinopathy. Features linked with diabetic retinopathy can be learned and used to train ResNet152V2. After the network is trained, it can be used to determine whether or not new fundus images contain diabetic retinopathy.

3. DenseNet201: In 2017, Huang et al. created a convolutional neural network (CNN) called DenseNet201. It's a 201-layer deep CNN that performs exceptionally well on image classification tasks. For DenseNet201 to function, it is necessary to tightly interconnect all of the layers. The network is then able to pick up more nuanced details from the pictures. DenseNet201 can be used to extract features from fundus pictures for the detection of diabetic retinopathy. DenseNet201 may be taught to recognize diagnostic markers for diabetic eye disease. After the network is trained, it can be used to determine whether or not new fundus images contain diabetic retinopathy.

4. InceptionV3: As part of the Google Inception project, Szegedy et al. introduced InceptionV3. It was made to be fast and accurate, thanks to a careful balancing of the two factors (model depth and computational efficiency). InceptionV3 is built on a novel framework that incorporates a number of "inception modules." These modules are the network's building pieces; their clever construction enables the model to record features at several scales and resolutions in a single layer. The Inception module is structured as follows, and it's at the heart of InceptionV3's success:

- Branch 1: Utilizes a 1x1 convolutional layer to capture local features.

- Branch 2: Comprises a 1x1 convolutional layer followed by a 3x3 convolutional layer. This combination captures slightly larger contextual information.

- Branch 3: Involves a 1x1 convolutional layer followed by a 5x5 convolutional layer. It captures more global features.

- Branch 4: Uses 3x3 max pooling followed by a 1x1 convolutional layer. This provides a form of down-sampling while retaining valuable features.

The InceptionV3 model effectively captures features at multiple dimensions, which is crucial for image analysis tasks such as the detection of diabetic retinopathy.

5. EfficientV2B3 and NasNetMobile: This study introduces an innovative hybrid approach that mixes the advantageous features of the EfficientV2B3 and NasNetMobile models. The EfficientV2B3 model is designed to enhance the model's capacity to acquire intricate features from images, thereby improving its performance in image-related tasks. On the other hand, the NasNetMobile model is specifically engineered to optimise computational resources, ensuring efficient utilisation of computing power while maintaining satisfactory model performance. Upon completion of the training process, the trained model can be effectively employed to classify novel fundus images based on the presence or absence of diabetic retinopathy. The utilisation of multiple models in conjunction with their computational efficiency can enhance accuracy in comparison to a single model. The aforementioned Model exhibits favourable attributes that render it a suitable option for implementation within clinical settings that prioritise the significance of both accuracy and efficiency. The integration of multiple models in the hybrid approach enables the hybrid model to attain a commendable level of accuracy and efficiency in the detection of diabetic retinopathy.

VIII. Potential Challenges

- The first challenge we encounter is the limited dataset available for this implementation. As we know, training a robust and accurate detection system heavily relies on a large and diverse dataset. However, in the case of diabetic retinopathy, obtaining labeled images can be a daunting task. To address this, we must emphasize the importance of data sharing and collaboration among healthcare institutions, researchers, and AI experts. By pooling our resources and creating centralized repositories, we can enhance the dataset's size and diversity, ultimately leading to more reliable detection models.

- The second hurdle we face is the issue of image quality. In the real world, medical images can be riddled with noise, orientation inconsistencies, and even corruption. This can significantly impact the performance of our detection algorithms. To mitigate this challenge, investing in image preprocessing techniques becomes crucial. By employing state-of-the-art denoising, image rotation, and restoration algorithms, we can enhance the quality of the input images, thus ensuring better and more accurate results.

- The third challenge revolves around the imbalance of the dataset. Diabetic retinopathy manifests in different stages, ranging from mild to severe. However, due to the nature of the disease, certain stages might be more prevalent in the dataset than others. This imbalance can lead to biased detection outcomes, where the algorithm may favor the more common stages and neglect the rarer ones. To tackle this, we must focus on data augmentation and balancing techniques. By generating synthetic data and oversampling the minority classes, we can create a more balanced dataset, ultimately leading to better performance and increased accuracy of the detection system.

- The fourth challenge was overfitting. Now, in basic sense, overfitting is when the model starts to stop remembering the patterns of the datasets and starts to focus completely on the training sets and cramming their features. Our train data at first was getting about 98% train accuracy with an average of 76% test accuracy which indicates a low level of overfitting which we needed to solve using Dropout Layers and Regulizers.

- The fifth challenge was we were getting a low amount of accuracy and very high amount of computation time when using data augmentation. We used Keras library "ImageDataGenerator" which creates data augmented images using given parameters.

IX. Result

Model Name	Train Accuracy	Test Accuracy
Sequential	86.88%	80%
ResNet152V2	98.28%	81%
InceptionV3	97.03%	81%
Xception	98.04%	81%
DenseNet201	98.62%	84%
EfficientNetV2B3	96.49%	76%
NasNetMobile	98.58%	80%
EfficientNetV2B3 + NasNetMobile	96.36%	82%

Our main goal as of now is to add two latest and least used models and add them together to create a hybrid model with an acceptable rating. From the table it can be seen that EfficientNetV2B3 only achieved around 76% of test accuracy which is the lowest among the models we have tested our project. But to make it more acceptable we combined it with NasNetMobile with Ensemble blending method which also has second lowest test accuracy and added them together to increase the test accuracy to 82% which is the second highest only to be lower than DenseNet201 which has a test accuracy of 84%. One other observation is that we Sequential model was the fastest among all the models used with a very good accuracy.

X. Conclusion

Our approach to the implementation and analysis of diabetic retinopathy detection using pre-trained models from Keras using 7 models has shown interesting findings and results. Our journey has been fueled by the recognition that a successful deep learning implementation requires not only cutting-edge models but also a high-quality dataset, which we were fortunate to obtain through a Kaggle competition. As we delve into the world of diabetic retinopathy detection, it's important to acknowledge the critical role of Convolutional Neural Networks. CNNs models have proven to be a game-changer in image-based tasks, demonstrating exceptional ability in feature extraction and pattern recognition. Among the powerful CNNs, we chose a few models to see the data for its efficiency and superior performance. The hybrid model architecture strikes a balance between model size and accuracy, enabling us to achieve impressive results without an overwhelming computational burden.

Bibliography

1. Shekar, S., Satpute, N., Gupta, A. (2021, November 29). Review on diabetic retinopathy with deep learning methods. SPIE Digital Library. https://doi.org/10.1117/1.JMI.8.6.06090

2. Dai, L., Wu, L., Li, H., Cai, C., Wu, Q., Kong, H., Liu, R., Wang, X., Hou, X., Liu, Y., Long, X., Wen, Y., Lu, L., Shen, Y., Chen, Y., Shen, D., Yang, X., Zou, H., Sheng, B., Jia, W. (2021, May 28). A deep learning system for detecting diabetic retinopathy across the disease spectrum - Nature Communications. Nature. https://doi.org/10.1038/s41467-021-23458-5

3. Yi, S. L., Yang, X. L., Wang, T. W., She, F. R., Xiong, X., He, J. F. (2021, November 22). Diabetic Retinopathy Diagnosis Based on RA-EfficientNet. MDPI. https://doi.org/10.3390/app112211035

4. Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture. (n.d.). Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture — IEEE Journals Magazine — IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/9409084

5. Basu, S., Mukherjee, S., Bhattacharya, A., Sen, A. (2022, July 9). Segmentation of Blood Vessels, Optic Disc Localization, Detection of Exudates and Diabetic Retinopathy Diagnosis From Digital Fundus Images. arXiv.org. Retrieved August 17, 2023, from https://arxiv.org/abs/2207.04345v1

6. Gothane, Raju, Bhaskar, Divya. (2022, January). Diabetic Retinopathy Detection UsingDeep Learning. researchgate. Retrieved August 17, 2023, from https://www.researchgate.net/publication/361780805

7. Population Therapeutics and Clinical Pharmacology, J. O., Pradeep, J., Jeffery, N. E., Saranraj, M., Hussain, J. N. (2023, May 18). Journal of Population Therapeutics and Clinical Pharmacology. Enhanced Recognition system for Diabetic Retinopathy using Machine Learning with Deep Learning Approach — Journal of Population Therapeutics and Clinical Pharmacology. Retrieved August 17, 2023, from https://jptcp.com/index.php/jptcp/article/view/1903