DIABETIC RETINOPATHY DETECTION – A Comparative Analysis of VGG-19 & ResNet-50.

Radhika Sanap[1], Zishaan Sayyed[1], Rukmani Sivaramakrishnan[1], Abhishek Pillai[1]
Prof. Manoj Patil[2]
Department of Computer Engineering, Rajiv Gandhi Institute of Technology, University of Mumbai.
[1]radhikasanap24@gmail.com, [1]zishaansayyed@gmail.com, [1]rukmaniiyerrrr@gmail.com, [1]itsabhishekpillaii@gmail.com, [2]manoj.patil@mctrgit.ac.in

***Abstract*— Diabetic retinopathy is a leading cause of blindness globally, making early detection and accurate diagnosis crucial for effective treatment. Our project suggests the utilization of two deep learning models, namely VGG-19 and ResNet, to categorize diabetic retinopathy into four distinct categories, namely: No DR, mild DR, medium DR, severe DR, and proliferative DR. We conducted training and assessment of the models using a retinal image dataset available to the public and were able to achieve a high level of accuracy and consistent performance. Our comparative analytical results of the two models demonstrate the potential of deep learning models in improving the efficiency and accuracy of diabetic retinopathy diagnosis, which can help clinicians make informed decisions and provide timely treatment to patients.**

Keywords—Diabetic Retinopathy, VGG-19, ResNet, DR, NODR.

I. INTRODUCTION

Diabetic retinopathy (DR) is a prevalent complication of diabetes and is recognized as the primary cause of blindness among adults of working age globally. DR is a consequence of harm inflicted on the blood vessels in the retina due to prolonged exposure to elevated glucose levels in the bloodstream. The damage leads to blood and fluid leakage into the retina, as well as the growth of abnormal blood vessels. If left untreated, DR can lead to severe vision loss and blindness.

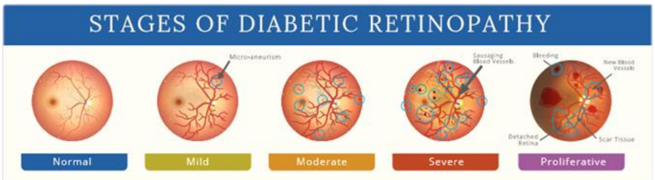

Fig. 1. Stages of Diabetic Retinopathy [1].

The severity of the disease is commonly categorised into five stages: No DR, mild DR, medium DR, severe DR, and proliferative DR. Non-proliferative retinopathy characterizes the initial phases of diabetic retinopathy, in which there is harm inflicted on the blood vessels in the retina, but no anomalous growth of blood vessels occurs. Proliferative retinopathy, on the other hand, is the advanced stage of the disease when abnormal blood vessels grow in the retina. Detecting and treating diabetic retinopathy (DR) at an early stage is crucial to prevent the onset of vision loss and blindness. The current standard method for detecting DR is a dilated eye exam, which is a time-consuming process that requires trained personnel. Therefore, there is a need for automated and accurate methods to detect and classify DR. Our research paper presents a machine-learning technique to categorize diabetic retinopathy (DR) into five stages utilizing retinal images. We utilize a convolutional neural network (CNN) to extract features from the retinal images automatically and classify them into the respective stages of DR. This approach has the potential to provide a more efficient and accurate method for DR detection, leading to earlier diagnosis and treatment of the disease and ultimately improving patient outcomes. In order to assess the efficiency of our suggested technique, we conduct a comparative study of two extensively applied deep learning models, namely VGG-19 and ResNet-50, for the classification of diabetic retinopathy (DR) using retinal images. The study uses the APTOS dataset, which has been extensively used in research studies for DR detection and classification. This dataset provides a diverse range of retinal images that enable the development of robust models that can generalize well to new data.

This research paper consists of five sections. Section II provides a literature survey of previous work related to the detection and classification of DR using deep learning techniques. Section III discusses the Kaggle Diabetic Retinopathy Detection Challenge dataset, which contains many retinal images with labels indicating the severity of DR, and the APTOS dataset used in this study. Section IV describes the methodology used, including the pre-processing of images and the training of the VGG-19 and ResNet-50 models on the dataset. Section V presents the results of the study, including a discussion of the performance of both models and a comparative analysis of their prediction capabilities. Finally, section VI provides concluding remarks and suggests potential directions for future research.

II. LITERATURE SURVEY

In paper [3], the system utilizes neural network and SVM models for the automatic detection and recognition of Diabetic Retinopathy, comparing patient images with normal images to detect the disease in its early stages. The system provides performance graphs for sensitivity, specificity, and accuracy, and includes a new concept for detecting Age-related Macular Degeneration. Our project attained an accuracy of 86.5% with a sensitivity of 90% and specificity of 87.5%. There is potential for further advancement by incorporating more extensive datasets and advanced classification techniques. In reference to [4], the utilization of retinal images for detecting diabetic retinopathy is suggested, as the process of manual diagnosis can be both resource-intensive and time-consuming. Identification of harm inflicted on the minuscule blood vessels in the retina is indicative of the presence of the disease. The proposed approach uses a neural network for classification, which is found to outperform the SVM model by 5%. In paper [5], the authors propose an unsupervised deep learning technique for early diagnosis of Diabetic Retinopathy (DR) using features extracted from the last convolution layer of an unsupervised ResNet50 model. The incorporation of rough set theory and fuzzy set theory has resulted in a significant improvement in the performance of the model, leading to state-of-the-art outcomes in the diagnosis of diabetic retinopathy. The integration of these techniques has enabled the system to attain exceptional results without any human intervention. The suggested model surpasses existing unsupervised algorithms, achieving an overall accuracy of 88.7%. Future work includes reducing the number of trainable parameters or making them adaptive to consume less memory. As per the research article [6], a DenseNet architecture-based deep learning technique is employed for the detection of diabetic retinopathy (DR) stages in high-resolution fundus images. On the APTOS dataset comprising of 3662 train images, the model was able to attain an accuracy of 0.9611, along with a quadratic weighted kappa score of 0.8981. The study also examined the performance of two CNN architectures for DR detection, DenseNet121 and VGG16. The authors of the research [7] developed a method for diabetic retinopathy diagnosis using a GPU trained on a set of pictures, with an accuracy of 81%.

.

III. DATASET

The Aptos dataset [8] is a publicly available dataset comprising 4,997 retinal images, which are divided into a training set of 3,662 photos and a validation set of 366 images. These images were rated on a scale of 0 to 4 by ophthalmologists based on the level of diabetic retinopathy present. Furthermore, the APTOS 2019 BD dataset used in the Asia Pacific Tele-Ophthalmology Society 2019 Blindness Detection experiment contains 3662 samples obtained from rural India. These samples were collected over an extended period and in diverse settings and environments by the Aravind Eye Hospital in India. Trained physicians analyzed and categorized the samples according to the International Clinical Diabetic Retinopathy Disease Severity Scale (ICDRSS) [9], which divides the samples into five categories, including no diabetic retinopathy, mild retinopathy, moderate retinopathy, severe retinopathy, and proliferative retinopathy. For the research, 3000 retinal fundus images were utilized from the APTOS (Asia Pacific Tele-Ophthalmology Society) database. The dataset was divided into 80:20 training, testing, and validation sets. The training set included 1000 pictures, 800 of which were classified as diabetic retinopathy (DR) and 200 as non-diabetic retinopathy (NODR). The testing set included 400 photos, 200 of which were diagnosed with DR and 200 with NODR. The validation set included 1000 pictures, 500 of which were diagnosed with DR and 500 with NODR. The dataset was balanced overall, with an equal proportion of DR and NODR photos.

IV. METHODOLOGY

A. Pre-processing of data.

1) Gussian Blur and Circular Crop : To prepare the retinal fundus images for analysis, we performed a series of pre-processing techniques to enhance the quality and clarity of the images. The pre-processing techniques applied included greyscale conversion, Gaussian blur, and circular cropping. The first pre-processing step involved converting the original color retinal fundus images into greyscale images. This was done to eliminate the color variations in the images and enhance the contrast between different retinal structures[10]. Greyscale conversion helped in reducing the dimensionality of the data, simplified image processing algorithms, and helped improved the performance of both the machine learning models that we've attempted to implement. The second pre-processing technique applied was Gaussian blur, which was used to smooth the images and reduce the noise and artifacts present in the original images. Gaussian blur helped to remove the high-frequency components of the image while preserving the low-frequency information[11]. This improved the overall quality of the images and enhanced the accuracy of the detection algorithm. The final pre-processing technique used was circular cropping, which involved extracting a circular region of interest from the retinal fundus images. This technique was used to eliminate the peripheral regions of the images that contained less relevant information and focus on the central region of the retina, where diabetic retinopathy typically manifested. The circular crop was centered on the optic disc, which is a prominent feature in the retina and can help in localizing the region of interest. In general, the use of pre-processing methods [12] is critical for enhancing the accuracy and reliability of diabetic retinopathy detection algorithms by enhancing the quality and consistency of retinal fundus images.

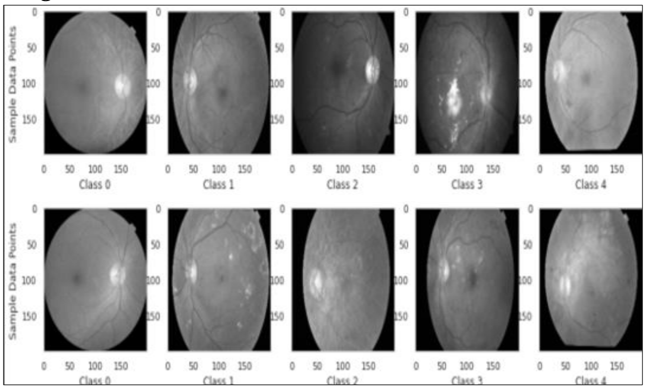

Fig. 2. Images after grey-scaling.

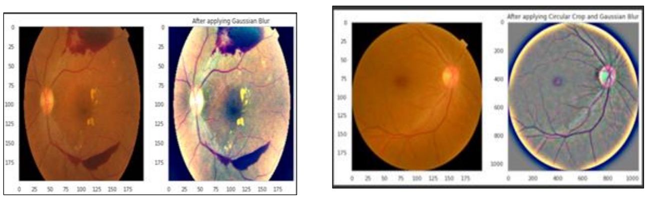

Fig. 3. Instance of images after applying gaussian filter and circular crop.

2) Data Augmentation: Data augmentation is a powerful tool for increasing the size of the dataset without collecting additional data, and it can be used in various machine learning and computer vision tasks to improve model performance. Our study on diabetic retinopathy utilized data augmentation methods to expand the training dataset and introduce diversity in the characteristics of the retina. By applying transformations such as rotation, scaling, flipping, and changing brightness and contrast, wegenerated new samples that were similar to the original data but had some degree of variation [13]. This helped in reducing the data size by 600 and improved the performance of the machine learning model by reducing overfitting and improving generalization.

B. *VGG-19 CNN*

After examining various models for detecting diabetic retinopathy, we discovered that VGG-19, a pre-trained convolutional neural network (CNN), is especially effective for this task. Originally developed for image classification tasks, VGG-19 has 19 layers, including 16 convolutional layers and 3 dense layers. It has demonstrated remarkable performance in numerous image recognition competitions [14]. Created by the Visual Geometry Group at the University of Oxford, VGG-19 is a deep CNN architecture that has become a popular pre-trained model, achieving state-of-the-art outcomes in several image recognition challenges. The network's design is based on the idea of using small 3x3 convolutional filters throughout the architecture, which leads to a better representation of spatial relationships between image features [15].

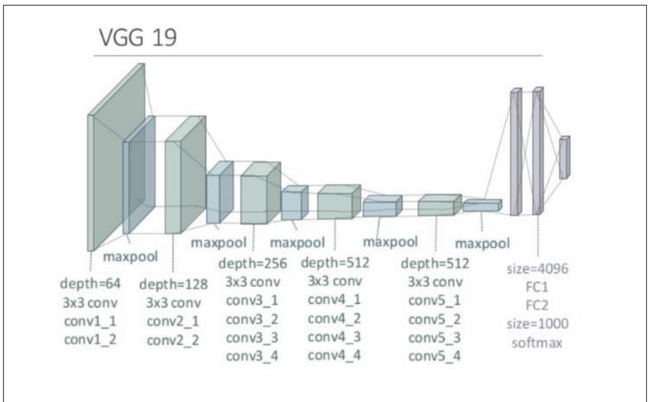

Fig. 4. VGG-19 Model [16]

To adapt VGG-19 for diabetic retinopathy detection, a new Keras model called "model" that is based on the pre-trained VGG-19 model called "vgg19_model" was created. To do this, a Sequential() function to create an empty model was created so that layers could be added to that. Then, all the layers were iterated over in the vgg19_model, except for the last layer, which is a dense layer responsible for the final classification task. For each layer, a new layer was added to the "model" that had the same configuration as the corresponding layer in vgg19_model. By using transfer learning, we were able to take advantage of the excellent feature extraction abilities of VGG-19 while still fine-tuning the model to detect diabetic retinopathy. This approach provided several benefits for the model, including a significant reduction in the number of trainable parameters due to the use of pre-trained weights from VGG-19. As a result, the training process was quicker and less prone to overfitting. In the VGG, it can be noted that VGG19 utilizes a sequential modeling approach, where the layers are arranged in a linear fashion. In this project, a self-made layer is integrated into the pre-trained VGG19 model to fine-tune it. Regarding the comparison between the two models for diabetic retinopathy detection, this model serves as a basis for comparison.

For this case, we have frozen all layers of the VGG-19 model except for the last one, which was randomly initialized and trained on the dataset. This approach was used to fine-tune the model for detecting diabetic retinopathy. By doing so, the model was able to benefit from the highly representative image features extracted by the pre-trained weights of VGG-19 while requiring training of only a relatively small number of weights in the last layer [17]. This allowed for higher accuracy to be achieved on the task using a smaller amount of training data.

Layer (type)	Output Shape	Param #
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv4 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv4 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv4 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102764544
fc2 (Dense)	(None, 4096)	16781312
dense (Dense)	(None, 2)	8194

Total params: 139,578,434
Trainable params: 8,194
Non-trainable params: 139,570,240

Fig. 5. VGG-19 Model Architecture after Transfer Learning

C. Resnet-50 CNN

In the ResNet section, it can be noted that ResNet50 uses functional modeling, where the layers are connected through shortcut connections to bypass one or more layers [18]. This allows the network to learn more efficient representations of the input data and to address the vanishing gradient problem that can occur in deep neural networks. ResNet50 is a highly optimized architecture that does not require any additional layers to be added to the pre-trained model, making it more efficient and easier to use for a wide range of tasks compared to VGG19. In our project, we implemented ResNet50 and evaluated its performance against VGG19 in detecting diabetic retinopathy. The VGG-19 model performed better at classifying true negatives, while the ResNet-50 model performed better at classifying false positives.

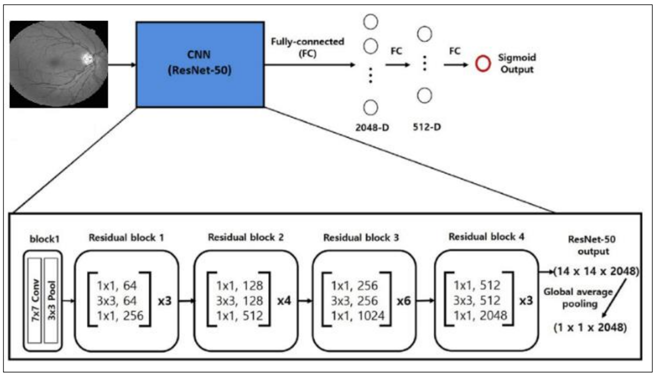

Fig. 6. ResNet-50 Model [19]

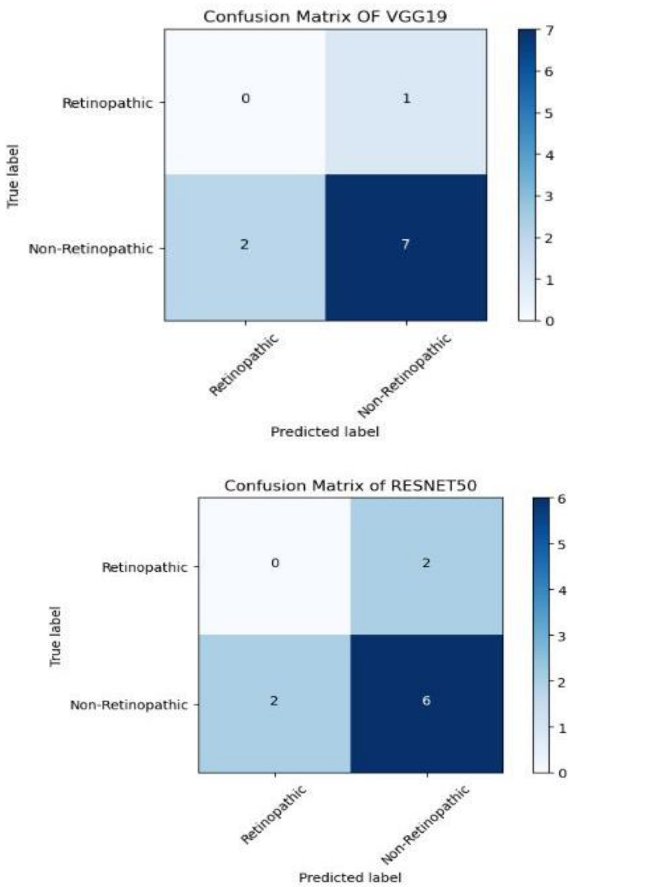

Fig. 7. Confusion Matrix for VGG-19 and ResNet Models.

V. Results and Analysis

A. Evaluation Metrics

1) Confusion Matrix: Confusion matrices were used to assess the classification performance of the VGG-19 and ResNet-50 models. To summarise the classification results of each model, two confusion matrices were produced. The first confusion matrix represents the VGG-19 model's classification performance, while the second confusion matrix shows the ResNet-50 model's classification performance. Each confusion matrix's rows and columns reflect the true and anticipated class labels, respectively. The VGG-19 model performed better at classifying true negatives, while the ResNet-50 model performed better at classifying false positives.

2) Accuracy Measures: In the study, the metrics of precision, recall, F1-score, and AUC-ROC were utilized. Precision was described as the proportion of correct positive predictions to the overall number of positive predictions.

precision = TP / (TP + FP) …(i)

where TP represents true positive predictions and FP represents false positive predictions.

To calculate recall, the study utilized the ratio of true positive predictions to the total number of actual positives. In other words, recall was defined as the proportion of correctly identified positive cases out of all actual positive cases:

recall = TP / (TP + FN) …(ii)

where FN represents false negative predictions. The F1-score was used as a combined measure of precision and recall, and was defined as:

F1-score = 2 * (precision * recall) / (precision + recall) …(iii)

The study also evaluated the AUC-ROC, which was computed by calculating the area under the receiver operating characteristic (ROC) curve. The ROC curve compares the true positive rate (TPR) against the false positive rate (FPR) at different thresholds. The AUC-ROC is a broad measure of the model's performance, with 1 representing ideal performance and 0.5 representing random chance.

3) Training and Loss Accuracy over time: The two popular deep neural network architectures, VGG-19 and Res -Net50, were evaluated on the dataset using accuracy and loss as evaluation metrics. In image classification tasks, accuracy is a metric that measures the percentage of correctly classified images. On the other hand, loss is a measure that quantifies the difference between predicted values and actual values. The results of the experiments were presented, as shown in Figure 8 and Figure 9, where the accuracy and loss over time for VGG-19 and ResNet50 were plotted. As seen in Figure 8, ResNet50 consistently outperformed VGG-19 in terms of accuracy, with an accuracy of 99% compared to VGG-19's accuracy of 97%. Additionally, as shown in Figure 8, ResNet50 had a lower average loss of 0.02 compared to VGG-19's average loss of 0.05. These results suggest that ResNet50 is a more effective architecture for the dataset, achieving higher accuracy and lower loss than VGG-19.

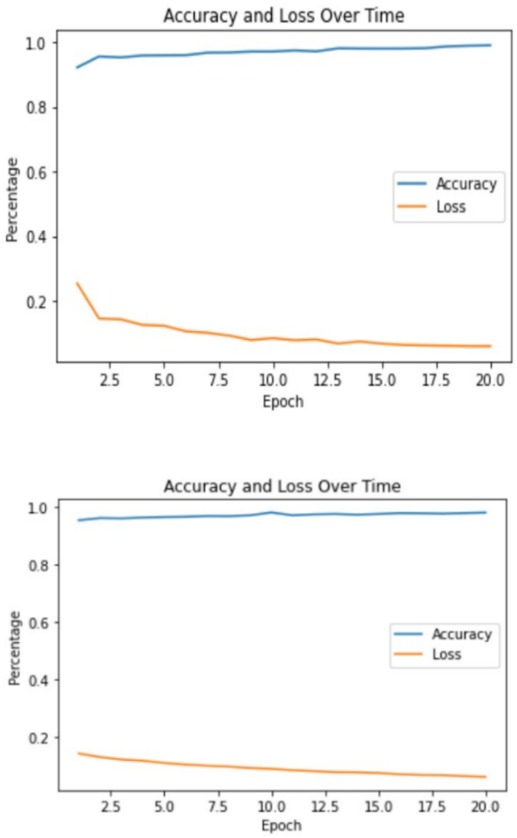

Fig.8. Accuracy and Loss Plot over Time for VGG-19 and ResNet-50.

B. *Comparative Analysis:* The study involved a comparison of two commonly used deep learning models, VGG-19 and ResNet-50, for a specific task. This comparison was made by analyzing the accuracy and loss metrics of both models over multiple epochs. The plot provides a comparative analysis of the learning dynamics of these models, which can offer valuable insights into their relative strengths and weaknesses.

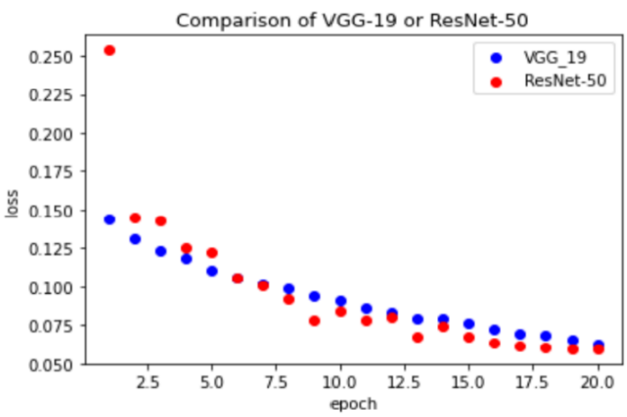

Fig. 9. Loss Scatterplot of VGG-19 and ResNet-50.

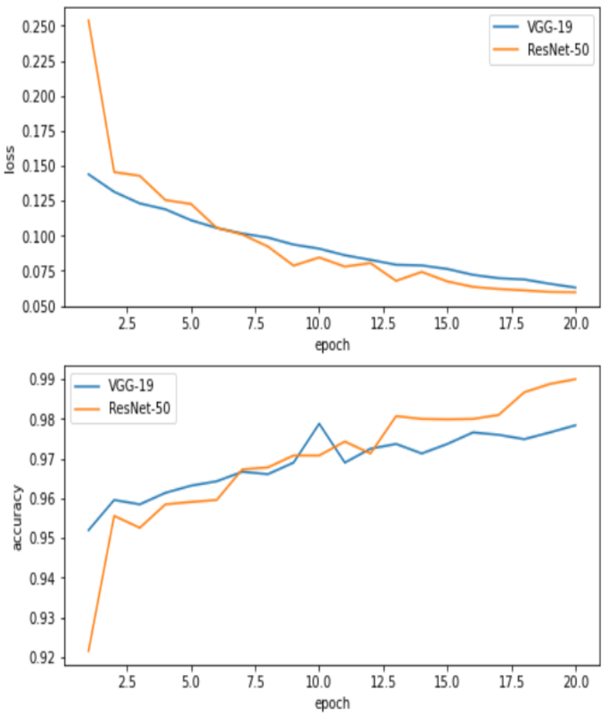

Fig. 10. Comparison plot of Accuracy and Loss for VGG-19 and ResNet-50.

Based on the analysis of the loss and accuracy plots for VGG-19 and ResNet-50, it can be concluded that ResNet-50 outperforms VGG-19 in terms of minimizing the loss and improving the accuracy of the model. The loss curve for ResNet-50 shows a steady decrease over time, indicating a more efficient learning process, while the loss curve for VGG-19 exhibits fluctuations and does not decrease as smoothly. Similarly, the accuracy curve for ResNet-50 shows a steady increase over time, whereas the accuracy curve for VGG-19 shows more variability and does not reach as high of maximum accuracy. Overall, these results suggest that ResNet-50 gives a more promising result over VGG-19.

VI. CONCLUSION

In this study, two deep learning models, VGG-19 and ResNet-50, were implemented for diabetic retinopathy detection. The web page developed for this purpose provides the results from both models for users to compare. The study conducted a comparative analysis of the two models to evaluate their effectiveness, and the findings revealed that ResNet-50 was more successful in terms of accuracy compared to VGG-19. The study's results emphasize the potential of deep learning models in the detection of diabetic retinopathy, underscoring the significance of selecting suitable models for such tasks. The study's future possibilities involve exploring alternative deep learning models and fine-tuning the parameters to enhance the detection accuracy. Moreover, further research could explore the potential of transfer learning approaches to utilize these models in other relevant medical image analysis tasks.

REFERENCES

[1] DIABETES & EYE EXAMS. (n.d.-c). https://cceyemd.com/diabetes-eye-exams/.

[2] Shukla UV, Tripathy K. Diabetic Retinopathy. 2023 Feb 22. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan–. PMID: 32809640.

[3] A. Gambhir, D. Sehrawat, H. Kumar, Y. Singh, and H. Saini, “Diabetic Retinopathy screening in human eyes using Image processing and segmentation,” *International Journal of Innovative Research in Computer Science & Technology*, vol. 8, no. 3, May 2020, doi: https://doi.org/10.21276/ijircst.2020.8.3.11.

[4] Yogesh Kumaram, Chandrashekhar M Patil, “A brief review of detection of Diabetic Retinopathy in human eyes using Pre-processing and Segmentation techniques”, *International Journal of Recent Technology and Engineering*, vol. 7, issue 4S2, December 2018.

[5] S. Rajkumar, R., and A. Grace Selvarani. “Diabetic Retinopathy Diagnosis Using ResNet with Fuzzy Rough C-Means Clustering.” *Computer Systems Science and Engineering*, vol. 42, no. 2, 2022, pp. 509–521, https://doi.org/10.32604/csse.2022.021909. Accessed 15 June 2022.

[6] S. Mishra, S. Hanchate and Z. Saquib, "Diabetic Retinopathy Detection using Deep Learning," 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), Bengaluru, India, 2020, pp. 515-520, doi: 10.1109/ICSTCEE49637.2020.9277506.

[7] S. Thorat, A. Chavan, P. Sawant, S. Kulkarni, N. Sisodiya and A. Kolapkar, "Diabetic Retinopathy Detection by means of Deep Learning," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2021, pp. 996-999, doi: 10.1109/ICICCS51141.2021.9432075.

[8] Tanwar, S. (2021, December 11). APTOS 2019 Blindness Detection - Towards Data Science. Medium. https://towardsdatascience.com/aptos-2019-blindness-detection-520ae2a4acc.

[9] Wilkinson, C. W., Ferris, F. L., Klein, R., Lee, P. P., Agardh, C., Davis, M. M., Dills, D., Kampik, A., Pararajasegaram, R., & Verdaguer, T. J. (2003, September 1). *Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales* [Video]. Ophthalmology; Elsevier BV. https://doi.org/10.1016/s0161-6420(03)00475-5

[10] Ping Zhang, Xiaoyou Shan and Feng Lu, "A novel eye detecting technology Based on Adaboost and gray-scale information," 2010 International Conference On Computer Design and Applications, Qinhuangdao, 2010, pp. V1-563-V1-566, doi: 10.1109/ICCDA.2010.5540684.

[11] E. S. Gedraite and M. Hadad, "Investigation on the effect of a Gaussian Blur in image filtering and segmentation," Proceedings ELMAR-2011, Zadar, Croatia, 2011, pp. 393-396.

[12] H. Poostchi, S. Khakmardan and H. Pourreza, "Diabetic Retinopathy dark lesion detection: Preprocessing phase," 2011 1st International eConference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran, 2011, pp. 177-182, doi: 10.1109/ICCKE.2011.6413347.

[13] K. Zhang, Z. Cao and J. Wu, "Circular Shift: An Effective Data Augmentation Method For Convolutional Neural Network On Image Classification," 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 2020, pp. 1676-1680, doi: 10.1109/ICIP40778.2020.9191303.

[14] K. Zhang, Z. Cao and J. Wu, "Circular Shift: An Effective Data Augmentation Method For Convolutional Neural Network On Image Classification," 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 2020, pp. 1676-1680, doi: 10.1109/ICIP40778.2020.9191303.

[15] M. Ali Sher, U. Muhammad, X. Yu and Q. Hu, "Fault Diagnosis of Rolling Element Bearing Using a Mesh of Continuous Wavelet Transform and Visual Geometry Group 19 (VGG-19)," 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China, 2021, pp. 102-106, doi: 10.1109/ICAICA52286.2021.9498027.

[16] Khaled Almezhghwi, Sertan Serte, "Improved Classification of White Blood Cells with the Generative Adversarial Network and Deep Convolutional Neural Network", *Computational Intelligence and Neuroscience*, vol. 2020, Article ID 6490479, 12 pages, 2020. https://doi.org/10.1155/2020/6490479.

[17] M. Mondal, M. F. Faruk, N. Raihan and P. Ahammed, "Deep Transfer Learning Based Multi-Class Brain Tumors Classification Using MRI Images," 2021 3rd International Conference on Electrical & Electronic Engineering (ICEEE), Rajshahi, Bangladesh, 2021, pp. 73-76, doi: 10.1109/ICEEE54059.2021.9719003.

[18] P. Patra and T. Singh, "Diabetic Retinopathy Detection using an Improved ResNet 50-InceptionV3 and hybrid DiabRetNet Structures," 2022 OITS International Conference on Information Technology (OCIT), Bhubaneswar, India, 2022, pp. 140-145, doi: 10.1109/OCIT56763.2022.00036.

[19] Kim, Mimi & Kim, Jong & Lee, Changhwan & Kang, Bo-Kyeong. (2021). Detection of pneumoperitoneum in the abdominal radiograph images using artificial neural networks. European Journal of Radiology Open. 8. 100316. 10.1016/j.ejro.2020.100316.

[20] Orozco Arias, Simon & Piña Duran, Johan & Tabares Soto, Reinel & Castillo, Luis & Guyot, Romain & Isaza, Gustavo. (2020). Measuring Performance Metrics of Machine Learning Algorithms for Detecting and Classifying Transposable Elements. Processes. 8. 1-19. 10.3390/pr8060638.