Proceedings of the 5th International Conference on Inventive Research in Computing Applications (ICIRCA 2023)
IEEE Xplore Part Number: CFP23N67-ART; ISBN: 979-8-3503-2142-5

Enhancing Diabetic Retinopathy Diagnosis with Inception v4: A Deep Learning Approach

R.Athilakshmi
Department of Computational Intelligence,
Faculty of Engineering and Technology,
SRM Institute of Science and Technology,
Kattankulathur-603203, Tamil Nadu, India.
athilakr@srmist.edu.in

R.Jansi
Department of Electronics and Communication Engineering,
Faculty of Engineering and Technology,
SRM Institute of Science and Technology,
Kattankulathur-603203, Tamil Nadu, India.
jansir@srmist.edu.in

Roshan Upadhyay
Department of Computational Intelligence,
Faculty of Engineering and Technology
School of Computing,
SRM Institute of Science and Technology,
Kattankulathur-603203, Tamil Nadu, India.
ru3578@srmist.edu.in

2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA) | 979-8-3503-2142-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICIRCA57980.2023.10220890

Abstract— **An automated system for the classification of Diabetic Retinopathy (DR) is introduced in this study. DR is a prominent cause of blindness globally, primarily impacting individuals aged 20 to 65. The proposed system can be an essential tool in the sustainable healthcare industry, providing timely and accurate diagnosis and treatment to patients with DR. Current manual screening by ophthalmologists can be time-consuming. Accurate categorization of DR is crucial to provide appropriate care and reduce the likelihood of permanent blindness. The use of big data and deep learning models in DR detection has the potential to revolutionize screening and diagnosis, leading to improved patient outcomes and reduced healthcare costs. This study compares traditional and CNN models, including Decision Trees, Random Forest, KNN, SVM, and Inception V4, to identify the best model for classifying retinal images into five stages of DR severity: No DR, Mild, Moderate, Severe, and Proliferative DR. Among the models tested, Inceptionv4 demonstrated the highest accuracy of 92.83% in classifying DR data. These findings demonstrate the potential of an automated system to improve DR diagnosis and management.**

Keywords— Diabetic Retinopathy, Inception V4, Deep learning, Computer vision, Sustainable Healthcare, Classification.

I. INTRODUCTION

Diabetes is a prevalent global disease that occurs when the body is unable to produce insulin, resulting in the inability to process glucose or blood sugar [1]. DR can occur as a result of diabetes. DR results in vision problems that can lead to permanent blindness. While DR initially causes mild vision problems or no symptoms at all, it can eventually lead to blindness, especially in individuals who have had diabetes for an extended period. Regular retina screenings are essential to identify and treat DR at an early stage to prevent blindness [2].

This study aims to automate the detection of Diabetic Retinopathy (DR) by classifying retina images into five stages of severity. This study has employed traditional models and convolutional neural networks (CNN) to automatically classify the retina images based on their characteristics. Our goal was to identify the best algorithm with the highest possible accuracy in detecting the retina images. The models used include Decision Trees, Random Forest, KNN, SVM, and Inception V4 [3, 4]. Experiments are conducted by using a publicly available Kaggle dataset containing 35,000 retinal images with 5-class labeling [5]. This study intends to establish a more effective technique for categorizing diabetic retinopathy for potential clinical benefits. Based on the existence of specific lesions, DR can be categorized into the five stages as shown in Fig. 1.

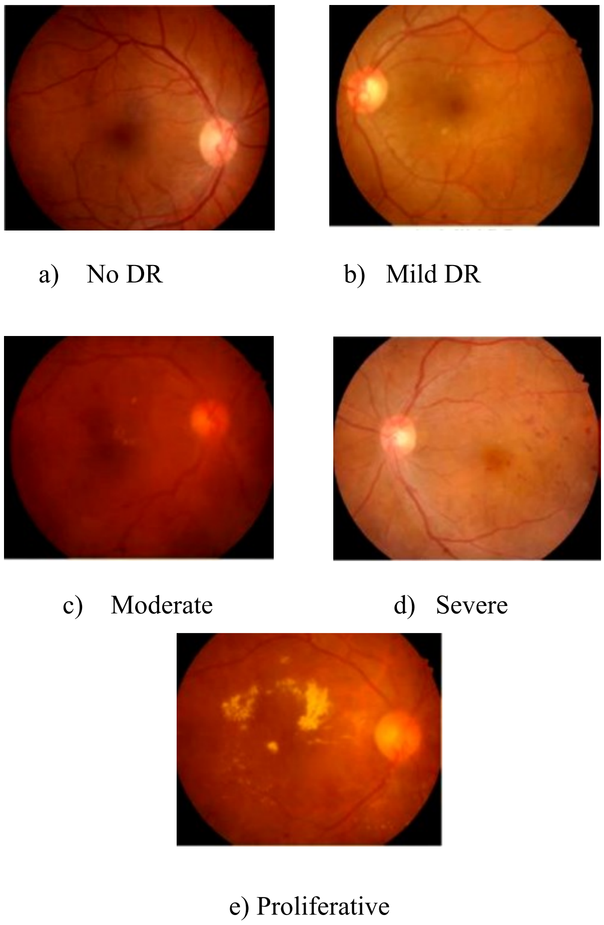

Fig. 1. Different stages of Diabetic Retinopathy

The motivation of this study lies in addressing the need for an automated system that can effectively classify

Diabetic Retinopathy (DR), a prevalent cause of blindness affecting individuals aged 20 to 65 worldwide. This study recognizes the importance of timely and accurate diagnosis in the sustainable healthcare industry, as manual screening by ophthalmologists can be time-consuming. The accurate categorization of DR is crucial for providing appropriate care and minimizing the risk of permanent blindness. By leveraging big data and deep learning models, the researchers aim to revolutionize the screening and diagnosis process, leading to better patient outcomes and reduced healthcare costs.

II. LITERATURE REVIEW

Numerous studies have been conducted to distinguish between patients with different stages of diabetic retinopathy (DR). In this survey [6], the researchers explored the foundational aspects of diabetes, encompassing its prevalence, complications, and the implementation of artificial intelligence (AI) methodologies for the early detection and classification of DR. Furthermore, the study investigated the utilization of AI techniques within this field. For example, transfer learning models have been utilized for automated DR detection in color fundus images, achieving an accuracy of 80% and 81% for DenseNet121 and NASNetMobile classification models, respectively [7]. In [8], the researchers provided compelling empirical data indicating that the utilization of residual connections resulted in a substantial acceleration of the training process for Inception networks. Several studies in the literature [9] have investigated the effectiveness of telemedicine-based DR screening. The findings consistently demonstrated that this approach was successful in detecting sight-threatening retinopathy and had a significant positive impact on compliance with annual retinal exams. These studies provided robust evidence supporting the efficacy of telemedicine-based DRS as a reliable method for identifying diabetic retinopathy and improving patient adherence to regular eye examinations. Other researchers proposed an approach that employed DenseNet-169 for classifying fundus images into five severity levels of DR [10]. Another study proposed a deep learning system called DeepDR, which utilized ResNet and Mask-RCNN and achieved an AUC of 0.934 for detecting different lesion types and classifying DR into various stages [11]. Through an extensive literature survey, it was observed that the DeepDR system exhibited exceptional sensitivity and specificity in the grading of diabetic retinopathy (DR). In addition to providing DR grading, the system also presented visual cues that aided users in identifying various lesion types and their respective locations. This scheme yielded enhanced diagnostic performance, aligning more closely with the cognitive approach employed by ophthalmologists. In addition, an automated classification system based on deep learning models was created to assign a severity grade for DR [12]. Here, an automated classification method for diabetic retinopathy (DR) severity grading was investigated. The study employed the utilization of convolutional neural networks, specifically VGG-16 and VGG-19, as machine learning models to analyze fundus images captured under different illumination conditions and sectors of vision. The objective was to generate a grading system that accurately assessed the seriousness of DR based on the analysis of these images.

This study intends to classify DR using the latest pre-trained deep learning model that is suitable for differentiating various stages of DR. We also discuss the pre-processing and regularization procedures necessary for tuning the hyper-parameters of the Inception v4 models.

III. PROPOSED METHODOLOGY

In the proposed methodology, the retinal images are first acquired. Then they are pre-processed and subjected to normalization. Finally, the images are classified using Inception V4. It is shown in Fig. 2. The selection of Inception V4 might be attributed to its ability to capture intricate features and patterns in the retinal images effectively. Inception V4 is a deep learning model that has been specifically designed for image classification tasks, incorporating various convolutional and pooling layers to extract rich spatial information. This architecture, combined with its training on large datasets, enables Inception V4 to learn complex representations and achieve high accuracy in image classification tasks.

A. Dataset description

The Kaggle dataset, made available by EyePacs [9], was utilized in this research.

B. Pre-processing

To reduce the computational complexity, all the images in the dataset were resized to 224×224. The intricate features of each image were retained in the resized image. Resizing of the images in the dataset to dimensions of 224x224 is typically achieved through a process known as image scaling. This involves adjusting the size of the images while preserving the aspect ratio and proportionality of the content within them.

C. Normalization

To standardize the pixel values of each channel in fundus images, normalization is performed such that the values fall between 0 and 1. Normalization of the pixel values in fundus images involves subtracting the minimum value and dividing by the range to ensure values fall between 0 and 1, ensuring consistency and comparability in analysis and processing.

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

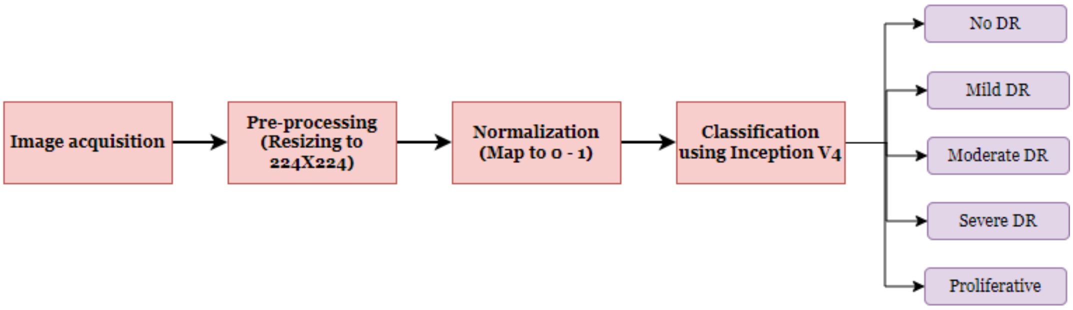

Fig. 2. Flowchart of the proposed classification scheme

This equation defines the normalization formula for fundus images with three channels, where y represents the normalized pixel value of the fundus image, x represents the input pixel value of the fundus image, x_{min} is the minimum pixel value, x_{max} and is the maximum pixel value.

D. Diabetic retinopathy classification

In this research, we have used the Inception V4 model architecture shown in Fig. 3 for classification. The Inception-v4 model, introduced by Google researchers in 2016, is a deep learning architecture. It is an improvement over its predecessor, Inception-v3 [14], and utilizes a number of advanced techniques. To process input images, the Inception-v4 model has been applied. The inception modules are a key feature of the model, which enables it to learn both local and global features in an efficient way. They do this by using a combination of filters of different sizes to deep information [15]. One of the main advantages of the Inception-v4 model is its high accuracy on image recognition tasks. The Inception-v4 model is also highly scalable.

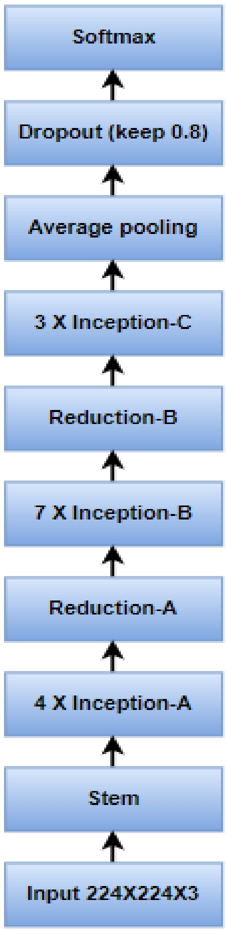

Fig. 3. Overall Architecture of Inception V4 model by Szegedy et al. [8]

The schema for the Inception-A, Inception-B, Inception-C, Reduction, and stem block of Inception-v4 is shown in Figures 4,5,6,7 and 8 respectively. One of the notable features of Inception V4 is the usage of Inception-ResNet architecture, which combines the advantages of both the Inception and ResNet architecture [13]. This combination results in a model that is not only efficient but also accurate, as it can extract features from images with varying levels of complexity.

In addition to its high accuracy, Inception V4 is also known for its speed and scalability, making it suitable for large-scale image recognition tasks. It has been trained on several large datasets, including the ImageNet dataset, and has attained good performance.

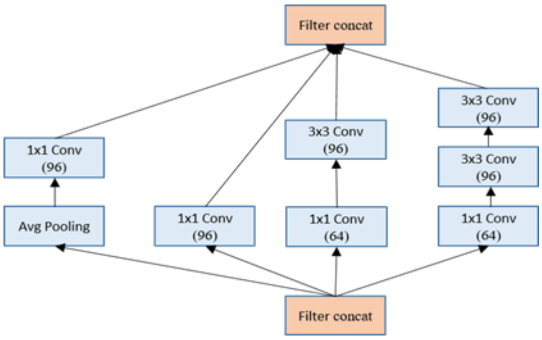

Fig. 4. Inception A Block as introduced in Szegedy et al.[8]

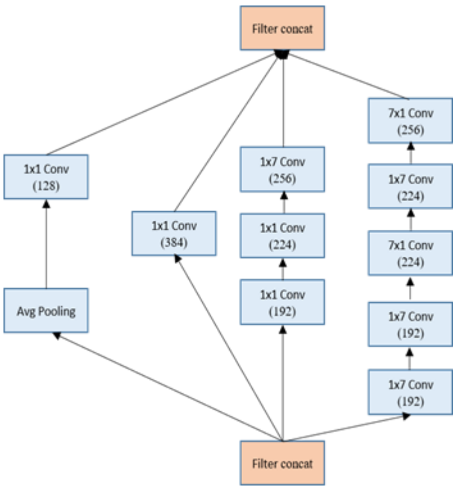

Fig. 5. Inception B Block as introduced in Szegedy et al. [8]

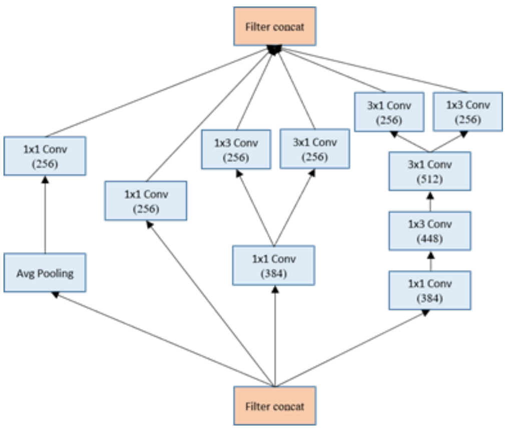

Fig. 6. Inception C Block as stated in Szegedy et al. [8]

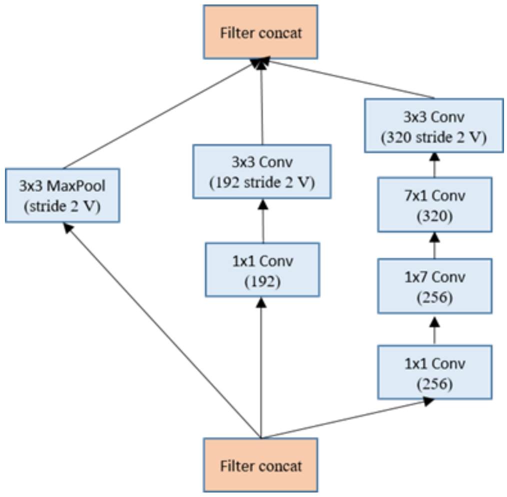

Fig. 7. The schema for Reduction module bySzegedy et al.[8]

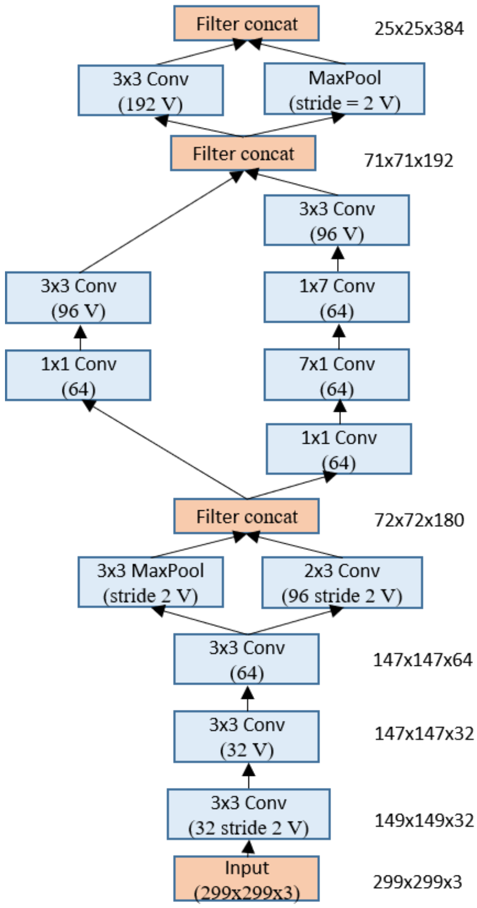

Fig. 8. Schema for Stem as stated in Szegedy et al. [8]

D. Working mechanism of Inception v4 architecture

The stem of the Inception v4 network receives the input image with dimensions of (299*299*3) and generates (35*35*384) convolutions as the output image. First, the output from the stem will be forwarded to four Inception-A blocks which factorize the incoming image into smaller convolutions by adding two ‘3*3’ convolutions in place of one ‘5*5’ convolution. As a result of this, the number of parameters will decrease by 28%. Next, the output image of the Inception-A block will move on to the Reduction block which alters the height and width of the grid, and the dimensions of the image were reduced from '35*35*384' to '17*17*1024'.

Then, the Inception 'B' block was applied to the output of the Reduction-A module that factorized the 7*7 convolutions into two asymmetric 1*7 and 7*1 convolutions and thereby reducing the number of parameters. Following that, the Reduction 'B' block was applied to the output image of the Inception'B' block and thereby downsizing the image to '8*8*1536' dimensions.

After that, three Inception 'C' blocks were applied to the downsized image of the Reduction 'B' block. Inception 'C' block factorizes the '3*3' image convolutions into two asymmetric convolutions of '1*3' and '3*1' and hence the total number of parameters goes down by 33%.

Finally, the output of the Inception 'C' block module was passed on to the average pooling, dropout, and softmax layer. Average pooling layer downsampled the feature map by finding the average value of each patch present in the grid. It also translates the image into a small amount by introducing a small piece of translation variance. The dropout layer helps the network to learn the data instead of memorizing the data and thus reduces the condition of overfitting. The final softmax layer converts the previous layer outputs into a vector of probabilities.

IV. RESULTS AND DISCUSSION

A. Dataset description

Our study employed a large DR dataset which was sourced from the Kaggle website (https://www.kaggle.com/c/diabetic-retinopathy-detection/data). For the development and evaluation of our model, we partitioned the dataset into two subsets. Specifically, 80% of the data was allocated for training purposes, while the remaining 20% was set aside for testing the performance of the trained model.

B. Hyperparameters

Table 1 presents the hyper-parameters utilized in our deep learning Inception V4 model, including a batch size of 16, an initial learning rate of 0.01, a momentum of 0.8, and a total of 25 epochs. The initial learning rate defines the step size at which the model updates its weights during training, and is set to 0.01 initially. Momentum refers to the factor by which the previous gradient update affects the current one, with a value of 0.8 indicating that the previous update has a significant influence on the current one. Finally, the number of epochs is set to 25 in this case.

TABLE I. HYPER-PARAMETERS

Parameters	Value
Batch size	16
Initial learning rate	0.01
Momentum	0.8
Number of epochs	25
Dropout	0.2

C. Hardware and software setup

This model was trained on an Intel Core i4 processor with 16GB RAM. Keras, a deep learning library, was utilized in conjunction with TensorFlow as the backend.

D. Quantitative analysis

The quantitative analysis of the Inception V4 model was done using parameters like accuracy and F1-score. Comparison of accuracy is shown in Table 2. This table presents the training and testing accuracy results for various models used in the study. The decision tree model achieved a training accuracy of 83.53% and a testing accuracy of 57.65%. The random forest model demonstrated higher performance with a training accuracy of 84.99% and a testing accuracy of 67.32%. The K-NN (K-Nearest Neighbors) model further improved the results, attaining a training accuracy of 87.84% and a testing accuracy of 73.53%. The SVM (Support Vector Machine) model showed even better performance, achieving a training accuracy of 91.75% and a testing accuracy of 75.41%. Finally, the Inception V4 model showcased the highest accuracy rates, with a training accuracy of 93.46% and an impressive testing accuracy of 92.83%.

TABLE II. PERFORMANCE EVALUATION OF DIFFERENT ALGORITHMS ON DR DATASET

Models	Training accuracy (%)	Testing accuracy (%)
Decision tree	83.53	57.65
Random forest	84.99	67.32
K-NN	87.84	73.53
SVM	91.75	75.41
Inception V4	93.46	92.83

Fig. 9 shows the comparison of the F1-score. The figure provides the F-scores of different classification algorithms used in the study. F-score is a metric that combines precision and recall to evaluate the overall performance of a classification model. According to the results, the decision tree algorithm obtained an F-score of 62.32, indicating its effectiveness in classification tasks. The random forest algorithm achieved a higher F-score of 75.92, suggesting improved performance compared to the decision tree. The K-NN (K-Nearest Neighbors) algorithm further improved the results with an F-score of 82.44, indicating better precision and recall. The SVM (Support Vector Machine) algorithm demonstrated even better performance with an F-score of 88.42, showing high accuracy in classification. Finally, the Inception V4 algorithm achieved the highest F-score of 92.74, indicating its exceptional performance in classifying the given data.

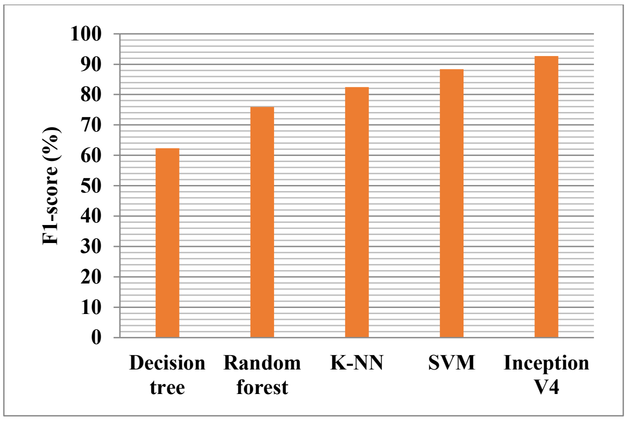

Fig. 9. Comparison of F1-score

V. CONCLUSION AND FUTURE WORK

The research paper described an automated system that can classify retinal images into five categories of diabetic retinopathy (DR) severity. To identify the best model with the highest accuracy for image classification, this study compared traditional models and pre-trained convolutional neural network (CNN) models. The study utilized several models, including Decision Trees, Random Forest, KNN, SVM, and Inception V4. It was observed that the F1-score for the Decision tree was 62.32%. Random forest achieved an F1-score of 75.92 and K-NN attained an F1-score of 82.44%. SVM and Inception V4 achieved an F1-score of 88.42% and 92.74% respectively. It was found that the latest version of the Inception models outperforms the existing methods in classifying DR data. Recently, interest has risen in the field of medicine for integrating deep learning-based assessment for the treatment of DR problems. The above research helps ophthalmologists to classify DR image data and to treat patients more effectively and efficiently. In future, investigating ensemble methods, such as combining predictions from multiple models, could potentially improve the overall classification performance. Ensemble techniques like model averaging or stacking could be explored to achieve better accuracy and robustness.

REFERENCES

[1] K. Göğebakan and M. Şah, "A Review of Recent Advances for Preventing, Diagnosis and Treatment of Diabetes Mellitus using Semantic Web," *2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Ankara, Turkey, 2021, pp. 1-6, doi: 10.1109/HORA52670.2021.9461282.

[2] Wong, T., Cheung, C., Larsen, M. *et al.* Diabetic retinopathy. *Nat Rev Dis Primers* **2**, 16012 (2016). https://doi.org/10.1038/nrdp.2016.12.

[3] Grading Diabetic Retinopathy from Stereoscopic Color Fundus Photographs—An Extension of the Modified Airlie House Classification: ETDRS Report Number 10, *Ophthalmology*, Volume 98, Issue 5, Supplement, 1991, pp. 786-806, ISSN 0161-6420, doi:10.1016/S0161-6420(13)38012-9.

[4] W. Chen, B. Yang, J. Li, and J. Wang, "An Approach to Detecting Diabetic Retinopathy Based on Integrated Shallow Convolutional Neural Networks," in *IEEE Access*, vol. 8, pp. 178552-178562, 2020, doi: 10.1109/ACCESS.2020.3027794.

[5] Chui, T.Y.P., Mo, S., Krawitz, B. *et al.* Human retinal microvascular imaging using adaptive optics scanning light ophthalmoscopy. *Int J Retin Vitr* **2**, 11 (2016), doi : 10.1186/s40942-016-0037-8.

[6] Bidwai P, Gite S, Pahuja K, Kotecha K. A Systematic Literature Review on Diabetic Retinopathy Using an Artificial Intelligence Approach. *Big Data and Cognitive Computing*. 2022; 6(4):152, doi: 10.3390/bdcc6040152.

[7] S. Sanjana, N. S. Shadin and M. Farzana, "Automated Diabetic Retinopathy Detection Using Transfer Learning Models," *2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, Dhaka, Bangladesh, 2021, pp. 1-6, doi: 10.1109/ICEEICT53905.2021.9667793.

[8] Szegedy, Christian, et al. "Inception-v4, inception-resnet and the impact of residual connections on learning." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 31. No. 1. 2017.

[9] Cuadros J, Bresnick G. EyePACS: an adaptable telemedicine system for diabetic retinopathy screening. J Diabetes Sci Technol. 2009 May 1;3(3):509-16. doi: 10.1177/193229680900300315. PMID: 20144289; PMCID: PMC2769884.

[10] 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.

[11] Dai, L., Wu, L., Li, H. *et al.* A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nat Commun* **12**, 3242 (2021). https://doi.org/10.1038/s41467-021-23458-5

[12] M. M. Krishnan, S. T. Ramya, K. Kirubanathavalli, S. Lalitha, J. Diofrin and M. Vadivel, "Deep Learning Approaches for Detecting Diabetic Retinopathy using CNN Models," *2022 International Conference on Automation, Computing and Renewable Systems (ICACRS)*, Pudukkottai, India, 2022, pp. 1096-1102, doi: 10.1109/ICACRS55517.2022.10029234.

[13] Das, D., Biswas, S.K. and Bandyopadhyay, S., 2022. Detection of Diabetic Retinopathy using Convolutional Neural Networks for Feature Extraction and Classification (DRFEC). *Multimedia Tools and Applications*,2022, pp.1-59.

[14] Alwakid, G., Gouda, W. and Humayun, M., 2023, March. Deep Learning-based prediction of Diabetic Retinopathy using CLAHE and ESRGAN for Enhancement. In *Healthcare MDPI*, Vol. 11, No. 6, p. 863.

[15] Oulhadj M, Riffi J, Chaimae K, Mahraz AM, Ahmed B, Yahyaouy A, Fouad C, Meriem A, Idriss BA, Tairi H. Diabetic retinopathy prediction based on deep learning and deformable registration. *Multimedia Tools and Applications*. 2022 Vol. No. 81, pp. 28709-27.