International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.044 ISSN: 2581-4621

Diabetic Retinopathy Detection Through Deep Learning Using CNN Wide-Net-X architecture

T. Kavitha[1], Basude Rohith[2], Akul Lakha[2], Akshay Chelikani[2]
[1]*Assocate Professor, Dept. of Electronics and Communication Engineering, Maturi Venkata Subba Rao (MVSR) Engineering College, Hyderabad, India.*
[2]*Student, Dept. of Electronics and Communication Engineering, Maturi Venkata Subba Rao (MVSR) Engineering College, Hyderabad, India.*

Abstract
In recent times, Diabetic Retinopathy (DR) has emerged as a critical complication for patients with diabetes, where the blood vessels in the retina are severely damaged, potentially leading to vision loss and, if left untreated, blindness. The World Health Organization has projected that by 2040, DR will impact around 224 million people. To address this issue, this research paper proposes CNN Wide-Net-X architecture model for image classification, which utilizes colour fundus images to detect Diabetic Retinopathy. The objective of this model is to enhance the accuracy and efficiency of the diagnostic process. For training and testing the model, the EyePACS dataset consisting of 5220 images is utilized, which is a widely accepted dataset for detecting Diabetic Retinopathy. To evaluate the performance of our model, we use metrics such as accuracy, precision, recall, and F1-score. The proposed CNN model is a significant step towards early detection and accurate diagnosis of DR. It is hoped that with the increased accuracy and efficiency provided by this model, patients with DR can receive timely treatment, thereby reducing the risk of vision loss and blindness.
Keywords— ***Diabetic Retinopathy, Deep Learning, Neural Networks, Fundus Images, Automated Detection.***

I. INTRODUCTION

Diabetic Retinopathy (DR) is a complication of diabetes that affects the retina and can cause vision loss in working-age individuals. The retina is the light-sensitive tissue located in the back of the eye that transmits visual signals to the brain. When diabetic individuals have high blood sugar levels, it can damage the small blood vessels in the retina, leading to DR. The disease can be classified into two main types: non-proliferative and proliferative. Non-proliferative diabetic retinopathy (NPDR) is the initial stage of DR and is characterized by microaneurysms, haemorrhages, and hard exudates. This stage can further be categorized into three types, namely mild, moderate, and severe NPDR. Proliferative diabetic retinopathy (PDR) is the most advanced stage of DR and is characterized by the growth of new blood vessels on the retina and in the vitreous humor. These newly formed blood vessels can be fragile and bleed, which can lead to severe vision loss and blindness. It is crucial to diagnose and treat DR early to prevent the development of advanced stages and avoid vision loss.
Convolutional Neural Networks (CNNs) have shown promise for automatic identification of Diabetic Retinopathy (DR) in fundus images. CNNs are effective in recognizing patterns in images and can accurately identify DR. CNNs are a type of deep learning algorithm that can recognize patterns in images, making them effective for tasks such as object recognition and image classification. These models have demonstrated high accuracy in identifying DR, and they can be integrated into telemedicine systems for remote screening. However, there is still room for improvement in terms of sensitivity and specificity. The purpose of this research is to look into the usage of CNNs for detecting DR in fundus images and to increase the model's performance by using different techniques. This study aims to investigate the use of different techniques to enhance the performance of CNNs in detecting DR, ultimately leading to more efficient and effective screening for this debilitating disease

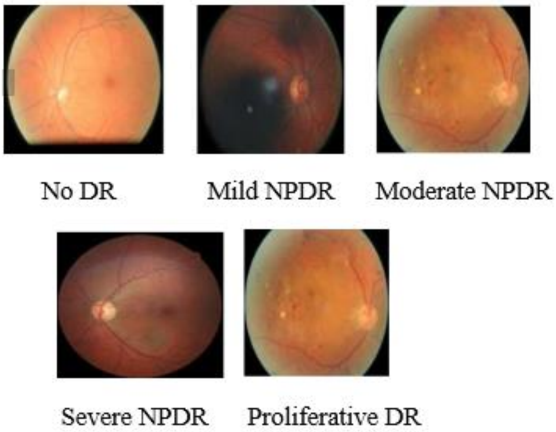

Fig. 1 Classification of DR

RELATED WORK

An intensive amount of research has been conducted and a great deal of effort has been put in to successfully perform image classification to classify color fundus images as either DR or non-DR based on both publicly available datasets and independently collected data. Depending on the datasets used and the training data labels, either binary classification or multi-class classification can be done. In the case of binary classification, we can only infer weather the retina in the colored fundus image has DR or not. In the case of multiclass classification, we can determine the severity of DR also. Some multiclass datasets have up to 5 classes. The methods used for DR detection range from Bayesian networks, tree algorithms to GANs, object detection methods and image classification methods using convolutional neural networks.

[1]Aziza et al proposed using a decision tree algorithm for detecting DR. They used the DRIVE and MESSIDOR databases. They have extracted the blood vessels followed by certain features of the blood vessels like area, perimeter etc. and used those features to build a decision tree using the CART algorithm. They have obtained sensitivity, specificity and accuracy of 91, 100 and 93 percent respectively.

[2]V.A.Aswale et al proposed a way to detect diabetic retinopathy by using support vector machines. For pre-processing they used CLAHE, vessel removal and illumination equalization. After this they performed feature segmentation and extraction. Walter, Flem, Zhang, Circular-Hough transform and Lazar were the feature segmentation and extraction algorithms used. They then performed classification of the fundus images using the SVM classifier on their own dataset of fundus images of the retina with their test set containing 30 images and obtained an accuracy of 93.33%.

[3]Md. Jahiruzzaman et al have used k-means clustering and fuzzy logic to detect diabetic retinopathy using the HRF dataset. In their approach they first perform eyeball edge detection using the sobel operator, followed by color compression using k-means. Then they classify the images using fuzzy classifiers. The sensitivity, specificity and accuracy obtained by them are 98.2, 89.8 and 92.3% respectively.

[4]Harry pratt et al proposed a custom convolutional neural network consisting of 10 convolutional layers, 2 fully connected layers and 1 softmax layer to classify DR images into 5 different classes on the EYEPACS dataset. The sensitivity, specificity and accuracy obtained are 30, 95 and 75% respectively.

The proposed model however has certain drawbacks such as having representational bottlenecks and overfitting. Also, a neural network can be trained faster with reduced dimensions. This can be done by factorizing filters of the first layers where there is a strong correlation among adjacent units without losing much information.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.044 ISSN: 2581-4621

II. PROPOSED APPROACH

Inception Inception Inception

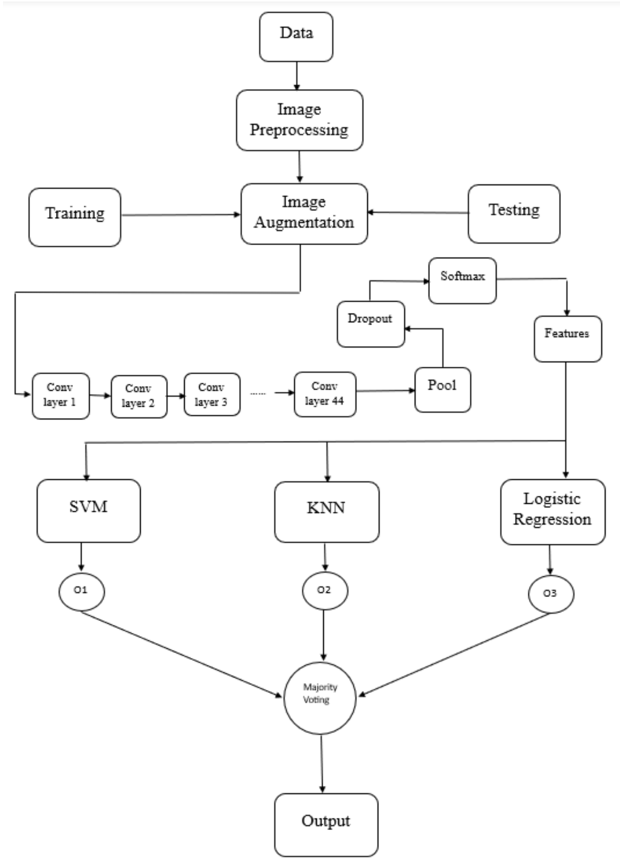

Fig. 2 Flowchart illustrating CNN Wide-Net-X

The proposed approach aims to improve upon existing CNN models used for DR detection by using a custom model based on the successful Inception architecture [5]. The approach seeks to enhance the predictive power of automated feature detection for classification into the five different DR severity classes, the idea or intuition utilized is that CNNs work best when multiple filters exist to capture various information. This point is conveyed more apparently in analysis of filter size effect on deep learning paper by Yunus Camgozlu et al [6] and In [7] paper by Nergis Tomen et al, A regular CNN employs the use of a single filter on a single layer. The use of a single filter might not successfully or completely capture all features present within a preliminary or abstract layer of any convolutional layer. On top of the existence of multiple filters in our model, there is another important aspect which deals with the sizes of the kernels. Here, different kernels or filters of different sizes are utilised, to eliminate the existence of a possibility where the 'perfect' kernel size to extract maximum information from the image is not chosen. The information extracted from each of these kernels of different sizes is then concatenated and passed on to the subsequent later. It involves the following steps:

A. Pre-processing

For the purpose of reducing the training time required for model and to increase the speed of inference, various pre-processing techniques are employed, including the resizing of all images to a fixed/constant size of 224x224.

Image Augmentation is performed using ImageDataGenerator class to reduce over-fitting, taking into consideration the relative size of each of the classes with respect to each other, to obtain a greater variety and diversity of images by using various techniques including adjusting the rotation_range, zoom_range, shear_range, horizontal_flip parameters or translating the image, along with the introduction of a little bit of noise to significantly improve the various generalization rules of the model.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.044 ISSN: 2581-4621

Next pixel brightness transformation is applied for enhancing the distinction of features from the background. We do this by applying the gamma correction method for each pixel after scaling the image pixel intensities to the range of 0.0 to 1.0. The computation can be formulated as follows:

$O = I \wedge (1 / G)$

Where I is the input image and G is gamma value

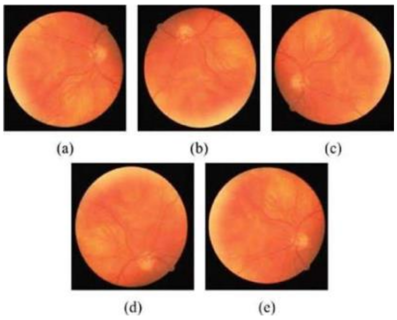

Fig. 3: First image (a) is original image and remaining are augmented images of the original image

B. Optimization technique

In the paper written by Amit Verma et al [8] they propose a method to reduce computational cost by manipulating filters. In our model we simply factorized certain filters such as 3*3 filters into 3*1 and 1*3 filters and 5*5 filters into a 3*3 and a 1*1 filter in order to reduce the computational cost as shown in figure - 4. By factorizing filters what we do is reduce the number of product operations for each convolutional operation. This also reduces the time it takes to train the model to a certain extent.

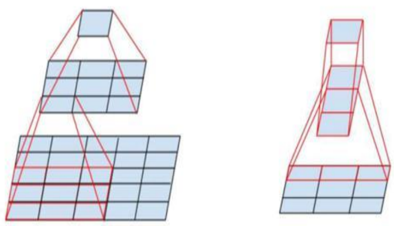

Fig. 4: Factorization of filters

C. Inception model

The model consists of 48 convolutional layers. The first 44 layers make use of the inception modules. Inception model consists of a mask of different sizes as discussed in optimized filters, pooling layer, Batch Normalization and Activation function. The activation function used for each of these layers is ReLu.

The last 4 layers are global average pooling, flatten and 2 dense layers. The activation function used by the dense layers are softmax and ReLu. The optimizer used is adam. The loss function we adopted was categorical cross entropy, which is perfect to utilise during a multi-class classification problem. The model is trained for 40 epochs on the set of augmented images and their labels.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.044 ISSN: 2581-4621

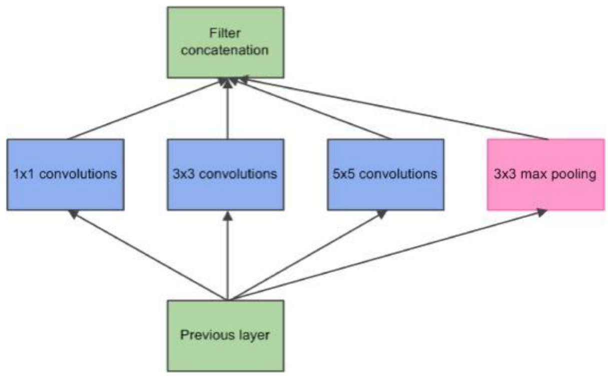

Fig. 5: Module of the inception layers used in the model

D. Classification

The classifier algorithms like SVM, decision trees and KNN are used for classification of images for better performance and accuracy which alone are not capable of learning complex behaviour. Majority voting is used for comparing the results from all algorithms and the algorithm with more accuracy is used for displaying the result.

III. RESULTS

This section elaborates the performance of the model. To evaluate the performance of our model we used 4 metrics. The metrics used are accuracy, precision, recall and f1-score. Each of them conveys a particular aspect of the performance of our model. Using multiple metrics gives us a comprehensive understanding of where the model is excelling or lacking.

The metrics used for evaluating the performance are as follows:

Accuracy: The fraction of correct predictions produced by a model out of all predictions made.

recision: The fraction of TP predictions made by a model out of all positive predictions.

Recall: The fraction of TP predictions made by a model out of all positive observations.

F1 Score: A precision-recall measure determined as the harmonic mean of precision and recall. It strikes a compromise between precision and recall, which is especially important when the classes are uneven.

A confusion matrix is a table that is often used to describe the performance of a classification model on a set of data for which the true values are known.

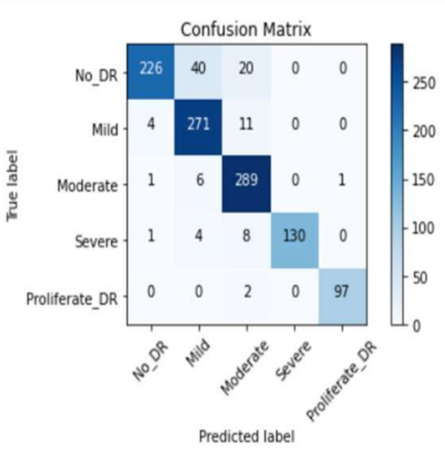

Fig. 6: Confusion matrix

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.7 May - June – 2023
DOI:10.46647/ijetms.2023.v07i03.044 ISSN: 2581-4621

The metrics described above are obtained from the values in a confusion matrix. Since our dataset has 5 classes i.e., No DR, Mild DR, Moderate DR, Severe DR, Proliferate DR, the confusion matrix will be a 5*5 matrix.
For the dataset all the metrics described above are calculated for 5 classes instead of the conventional positive and negative classes.
For the dataset, all the metrics described above are calculated for 5 classes i.e., 286 No-DR, 286 Mild DR, 297 Moderate DR, 143 Severe DR, and 99 Proliferate DR images. The metrics for these classes are then averaged to obtain consolidated metrics.
Thus, the use of classifier algorithms for detection of the DR at different stages improved the accuracy of detection. And majority voting at the end of the classifier algorithms helps to attain the maximum accuracy from the algorithms.
The trained model obtained an accuracy of 93% on the training dataset and an accuracy of 92.8% on the testing dataset. The precision, recall and F1-score obtained are 97%, 93% and 93% respectively.

DISCUSSION AND CONCLUSION

The main distinguishing metric is the sensitivity. Sensitivity is crucial in medical diagnosis as false negative predictions can cause delays in treatment.
Class imbalance can also affect the accuracy of the model, as is the case where most of the images used belong to a single class. This can lead to overfitting, resulting in false positives. The proposed architecture for improved accuracy is not the only reason for this, as the treatment of class imbalance is also a crucial factor. The confusion matrix shows that most false positives are attributed to the class with the highest number of images, indicating the need to address the imbalance in the dataset.
This paper may serve as a stepping stone for improvements in automated medical image diagnosis as well as image processing using deep learning in general. DR continues to grow as a major illness among diabetic patients and improved diagnosis is required as soon as possible. To be able to see faster progress in this arena we may need to overcome certain hurdles. One of these hurdles is the presence of only 2 publicly available datasets with more than 10,000 color fundus images of the retina. Better color fundus cameras may be used to capture better resolution images. Another issue is the slight uncertainty of the image labels found in the dataset. The labels are assigned after careful diagnosis by an ophthalmologist which are not 100 percent accurate. This may hinder deep learning models from being able to learn all the features of an image belonging to a specific class.
To conclude we have illustrated our approach for an improved, automated prediction of DR using a CNN Wide -Net-X architecture to the best of our abilities.

ACKNOWLEDGMENT

This study was carried out under the supervision and guidance of the T.Kavitha, Assoc Prof of Maturi Venkata Subba Rao (MVSR) Engineering College.

REFERENCES

[1] Aziza, Elaouaber Zineb; Mohamed El Amine, Lazouni; Mohamed, Messadi; Abdelhafid, Bessaid, “Decision tree CART algorithm for diabetic retinopathy classification” (2019). [IEEE 2019 6th International Conference on Image and Signal Processing and their Applications (ISPA) - Mostaganem, Algeria (2019.11.24-2019.11.25)]
[2] “Detection of microaneurysm in fundus retinal images using SVM classifier” V.A. Aswale, J.A. Shaikh, IJEDR, Volume 5, Issue 4, ISSN: 2321-9939
[3] “Detection and Classification of Diabetic Retinopathy Using K-Means Clustering and Fuzzy Logic”,Md. Jahiruzzaman, A. B. M. Aowlad Hossain, 18th International Conference on Computer and Information Technology (ICCIT), 2015.
[4] Pratt, Harry; Coenen, Frans; Broadbent, Deborah M; Harding, Simon P; Zheng, Yalin (2016). “Convolutional Neural Networks for Diabetic Retinopathy”. Procedia Computer Science, 90 (2016) 200 – 205.

[5] "Going deeper with convolutions", Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14).
[6] "Analysis of Filter Size Effect In Deep Learning" Yunus Camgözlü1, Yakup Kutlu, Journal of Artificial Intellicence with Application, 2020.
[7] "Spectral Leakage and Rethinking the Kernel Size in CNNs", Nergis Tomen, Jan C. van Gemert, ICCV paper, 2021.
[8] "Computational Cost Reduction of Convolution Neural Networks by Insignificant Filter Removal", Amit Verma, Toshanlal Meenpal, and Bibhudendra Acharya, Romanian journal of information science and technology Volume 25, Number 2, 2022.
[9] M. Chetoui, M. A. Akhloufi and M. Kardouchi, "Diabetic Retinopathy Detection Using Machine Learning and Texture Features," 2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE), Quebec City, QC, 2018, pp. 1-4.
[10]"Rethinking the Inception Architecture for Computer Vision", Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[11]E. V. Carrera, A. Gonz´alez and R. Carrera, "Automated detection of diabetic retinopathy using SVM," 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Cusco, 2017, pp. 1-4.
[12]M. Manjramkar, "Survey of Diabetic Retinopathy Screening Methods," 2018 2nd ICOEI.
[13]K. K. Palavalasa and B. Sambaturu, "Automatic Diabetic Retinopathy Detection Using Digital Image Processing," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, 2018, pp. 0072-0076.
[14]S. Kanth, A. Jaiswal and M. Kakkar, "Identification of different stages of Diabetic Retinopathy using artificial neural network," 2013 Sixth International Conference on Contemporary Computing (IC3), Noida, 2013, pp. 479-484.
[15]A. Herliana, T. Arifin, S. Susanti and A. B. Hikmah, "Feature Selection of Diabetic Retinopathy Disease Using Particle Swarm Optimization and Neural Network," 2018 6th International Conference on Cyber and IT Service Management (CITSM), Parapat, Indonesia, 2018, pp. 1-4.
[16]Mujahid M, Rustam F, Álvarez R, Luis Vidal Mazón J, Díez IT, Ashraf I. "Pneumonia Classification from X-ray Images with Inception-V3 and Convolutional Neural Network". Diagnostics (Basel). 2022 May 21;12(5):1280.
[17]Li, Yunghui & Aslam, Muhammad & Yang, Kai-Lin & Kao, Chung-An & teng, Shin-You. (2020). "Classification of Body Constitution Based on TCM Philosophy and Deep Learning". Symmetry. 12. 17. 10.3390/sym12050803.
[18]Boonyuen, Kitinan & Kaewprapha, Phisan & Weesakul, Uruya & Srivihok, Patchanok. (2019). "Convolutional Neural Network Inception-v3: A Machine Learning Approach for Leveling Short-Range Rainfall Forecast Model from Satellite Image".
[19]Szegedy, Christian & Ioffe, Sergey & Vanhoucke, Vincent & Alemi, Alexander. (2016). "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning". AAAI Conference on Artificial Intelligence. 31.
[20]Alyoubi, W. L., Shalash, W. M., & Abulkhair, M. F. (2020, January 1). Diabetic retinopathy detection through deep learning techniques: A review. *Informatics in Medicine Unlocked*. Elsevier Ltd. https://doi.org/10.1016/j.imu.2020.100377