Indian Journal of Computer Science and Technology
https://www.doi.org/10.59256/indjcst.20240302007
Volume 3, Issue2 (May-August 2024), PP: 60-64.
www.indjcst.com
ISSN No: 2583-5300

Automatic Diabetic Retinopathy Detection Using Resnet50 and Inceptionv3

Dr. K. Paramasivam[1], Jaspar vinitha sundari T[2], D. Sam Chrisvin[3]

[1]*Professor, Department of EEE, Kumaraguru College of Technology, Coimbatore, Tamilnadu, India.*
[2]*Assistant Professor, Department of ECE, Kumaraguru College of Technology, Coimbatore, Tamilnadu, India.*
[3]*Student, Department of ECE, Kumaraguru College of Technology, Coimbatore, Tamilnadu, India.*

To Cite this Article: *Dr. K. Paramasivam[1], Jaspar vinitha sundari T[2] D. Sam Chrisvin[3], "Automatic Diabetic Retinopathy Detection Using Resnet50 and Inceptionv3", Indian Journal of Computer Science and Technology, Volume 03, Issue 02 (May-August 2024), PP: 60-64.*

Abstract: *Diabetic retinopathy (DR) is an eye disease in diabetic patients and is the cause of blindness in the population. Because of the complexity of color fundus images DR classification by humans is challenging and is also an error prone task. Thus, this paper proposes automated DR classification to detect diabetic retinopathy in advance. Traditional machine learning algorithms like Logistic regression (60%) and Support Vector Machines (70%) have been employed to detect the severity of DR using information from retina images in this work. Since machine learning Algorithms provide low accuracy, deep learning algorithms have been used to improve the accuracy of the model. To develop a model from scratch is a complicated task, which can be solved using another technique called transfer learning. In this work, two CNN models are trained - Resnet50 and Inception V3 with pre-trained imageNet weights. ResNet50 is used in two ways, as a classifier, fine tuning and InceptionV3 without fine tuning. The CNN model achieved a good accuracy around (85%) using InceptionV3 without fine tuning. The dataset used was provided by Kaggle which are high resolution images of the retina which are proven to show good accuracy. Since machine learning Algorithms provide low accuracy, deep learning algorithms have been used to improve the accuracy of the model. To develop a model from scratch is a complicated task, which can be solved using another technique called transfer learning. In this work, two CNN models are trained - Resnet50 and Inception V3 with pre-trained imageNet weights. ResNet50 is used in two ways, as a classifier, fine tuning and InceptionV3 without fine tuning. The CNN model achieved a good accuracy around (85%) using InceptionV3 without fine tuning. The dataset used was provided by Kaggle which are high resolution images of the retina which are proven to show good accuracy.*

Key Words: *Diabetic Retinopathy; Support Vector Machine; CNN, ResNet 50; Inception V3;*

I.INTRODUCTION

Diabetic Retinopathy (DR) is one of the main factors affecting vision for people over a wide range of ages and has affected nearly 93 million individuals. DR can progress to a vision loss if not detected in early stages with perfect clinical screening. Several complications can be negated if the DR can be detected in the early stages. The early stages of DR can be visualized as micro-aneurysms and exudates in the retina. To avoid several complications such as vision loss, the next stage of having hemorrhages should be detected at the earliest. Severe Class of DR is characterized to have intraretinal hemorrhages and increased count of micro-aneurysms in various quadrants. Proliferative stages have been classified as one of the severe stages in DR which can be visualized from Neovascularization which could be present on disk or elsewhere. The current trend of detecting DR is by using retinal images and thereby evaluated manually by efficient professionals. The DR screening process has been tricky for manual readers which has increased inconsistency and reduced accuracy of results thereby holding a challenge in the traditional DR detection process. Thus, there exists a need for Automated grading services for increasing efficiency, regenerative capability and wide range of coverage in screening. Algorithms are used to learn the source data and thereby provide better and enhanced outcomes on DR screening. The algorithms are developed based on Retinal image study, which includes mathematical modeling and wavelet modification.

Some work has been done on various frameworks in machine learning and deep learning methods for early identification of DR. For example, research reported 8 3 stage framework has been implemented for automated grading of DR which applies intensity normalization, data balancing in the preprocessing stage of raw data. QIV-3 model9 is used while the images are resized to 600*600 pixels. Preprocessing along with OD Segmentation (Optical Disc) is done for converting the image into HSV plane and further those images are divided into 4 Quadrants9. Implementation of machine learning algorithms has been predominant in detecting DR especially with the stage of Microaneurysms. The extraction of 25*25 patch pixels extracted from baseline images in the Diabetic Retinopathy Database, Calibration Level 1, was used as input for classifiers such as Forest Randomization (RF), Neural Networks (NN) and Support Vector Machine (SVM) and extract features such as principal component analysis and random forest which enhanced efficiency (AUC) by 0.962 to 0.983 and also an F measurement of 0.913 to 0.926 is achieved1. Stack Sparse Autoencoder (SSAE) of DL instances in fundus images are used2, Here SSAE is fed to the classifier which is used to classify the image as MA or Non-MA. Here SSAE learned directly from raw images and distinguished them automatically and classify using the Soft Max classifier. After fine tuning F-measure is 0.913 and AUC is 96.2. Various Transfer learning methods have been implemented in DL models[3] where Classification of various aneurysms of the retina eyes with diabetic retinopathy is achieved. Models implemented on datasets on ImageNet, and a few layers which are present at first of the model have been frozen. The first few layers of the model. It requires a very accurate learning model to classify stages correctly. 95.9% by using LR, 91% by using SGD, 95.9% by using NN, and 93.4% by using HE and a high AUC of 99.7% is captured. Transfer learning methods[4] in both Machine Learning and Deep Learning models like SVM and CNN along with feature extraction are done. The Convolutional Neural Network (CNN) extracted high level features from the last connected layer during the transfer learning process and it is fed as input for support vector machines for the classification process. Using CNN with finetuning, computation time for classification is reduced. During transfer learning, extraction of features is done using Alex net, VGG net, InceptionNet, Google Net, and Densenet. A several classifications in Resnet 50 and VGG-16 has been made[5], Between the ResNet50 model and VGG-16 model in the DR Classification ResNet50 provided a top performance with accuracy of 70% where VGG-16 provided 25%, which shows that in images classification, specifically in many images, ResNet50 architecture is the best. Various Feature extraction methods are done with Resnet 50 has been fed as an input to Random Forest for Classification purposes[6]. An accuracy of 96% and 75.09% for the Messidor-2 dataset is obtained. Since Random Forest classifier is used instead of Restnet-50 classifier, improved accuracy of 96% is obtained[6]. Stages which include Hard exudates are identified and the parameters such as Sensitivity are detected [7] using Support Vector Machine (SVM). SVMs have detected DRNPs with almost 95% sensitivity, but DRNPs can be classified with an average accuracy of 85%. SVM consistently outperforms other machine learning algorithms. Pre-processing techniques such as intensity normalization and augmentation have been implemented[8] which shows the classification accuracy of ResNet of about 86.67%. CNN models such as InceptionV-3 have been implemented on Quadrant based Automated Grading Systems. Since it is also one of the predominant models for detection of various stages in DR. Inception V-3 is used[9] in which image enhancement and optical disc removal pipeline along with data augmentation stage has been done in the architecture. The system achieves 93.33% accuracy with a minimized cross-entropy loss of 0.29. Data Augmentation is one of the important processes for increasing the data volume thereby improving the measurement parameters. For Vessel Detection, Optical Disc Detection etc various preprocessing methods have been implemented[12] such as shade correction, Median Filter, Adaptive Contrast Enhancement. A Gaussian filtering model was used[14] for retinal blood vessel segmentation on the input fundus images. Data expansion steps[10] such as histogram equalization, optical disc position, and quadrant trimming to improved network performance have been implemented in a deep neural network named Inception Res net V-2. 8 Layered architectures such as Alex Net are implemented[11] for various feature extraction. The extracted features are sent to a pixel-by-pixel neural network (NN) classifier or support vector machine (SVM) classifier for automatic DR scoring. By obtaining an Alex Net based architecture and a pixel-wise NN classifier, the final system achieves accuracy, sensitivity, and specificity 95.73%, 95.73% and 98.51%. Implementation of CNN with Selective Sampling strategy (SeS) is done[13] which addresses common issues in medical image analysis tasks. The image level performance is also improved as 0.919 and 0.907 for the SeS and NSeS from the datasets.

The sections of the paper are structured in the following order. Section 3 explains the dataset description, preprocessing and algorithms used. Section 4 explains results and discussion. Section 5 explains the conclusion and future work.

II.MATERIAL AND METHODS

Datasets are obtained from Kaggle, which are high resolution images of the eye and various identification methods have been used to identify defects in the eye using each individual pixel, so that classification of images can be done. The following table represents the percentage of each of the classes present in the data. Obtained data is classified into 5 Classes based upon the severity of the DR in the eye.

Table no 1: Image data set classes

Class	Name	Number of images	Percentage
0	Normal	25810	73.48%
1	Mild NPDR	2443	6.96%
2	ModerateNPDR	5292	15.07%
3	SevereNPDR	873	2.48%

Data preprocessing:

Each image has a black background apart from the retina fundus which is not required for any classification algorithms. Hence there exists a need for a preprocessing method for removing the black background. Removal of black background is implemented which reduces the size of the image and makes processing faster. Adaptive Histogram Equalization (CLAHE) is used to make the blood vessels, exudates and hemorrhage more prominent. Resizing the image for the algorithm specifications $64 \times 64 \times 3$ pixels, $224 \times 224 \times 3$ pixels and $299 \times 299 \times 3$ pixels is done.

Algorithms used

Various Machine Learning and Deep Learning Approaches Logistic Regression, Support Vector Machines and Convolutional Neural Networks include ResNet50, Transfer Learning using ResNet50, InceptionV3 are used in this work to derive the Classification of Diabetic Retinopathy

III.RESULTS AND DISCUSSIONS

Preprocessing data

The images used were high resolution images with lots of information which are not needed for the algorithm. So, it is

necessary to preprocess the images to remove this information and enhance the features for better results.
Removing black border is done as follows. The retina fundus in the images were surrounded by a black border which must be removed because it is not useful to the algorithm.

Contrast Limited Adaptive Histogram Equalization (CLAHE) is used to differ from traditional histogram equalization in terms of computing various histograms, Adaptive histogram equalization is implemented to improve image contrast and definition at the edges. The Adaptive method measures and computes histograms, each on a range of distinct sections of the image. Thus, the distinctive sections are used to correct the lightness values of the image. Image Definition at the edges and local contrast are enhanced in a region, thereby adaptive methods are appropriate for histogram equalizations. CLAHE helped to make the exudates, hemorrhages, blood vessels and micro-aneurysms more prominent. According to the requirement of the algorithm the image was resized to 64 × 64, 224 × 224 or 299 × 299 pixels

Fig. 1. Stages of preprocessing, Original Image, Image after removing black border and Image after CLACHE.

Logistic Regression

First, simple logistic regression is applied over preprocessed 64×64×3 images. Here a trivial method is used to train the logistic regression model over the intensity values of images for each pixel. The 64×64×3-pixel images are reshaped to a flat 12288 (64 × 64 × 3) vector. Using this as a feature vector for each image, the classifier was trained. Though the accuracy was (around 60%) but this worked as a benchmark for the rest of our classifiers.

Support Vector Machine

The same approach as in Logistic Regression is used in Support Vector Machines (SVM) . Here, the accuracy is improved (around 70%) as SVM would have been able to fit a better hyperplane than expected. SVM with default parameters and various tuning parameters techniques were used but it didn't help much.

ResNet50

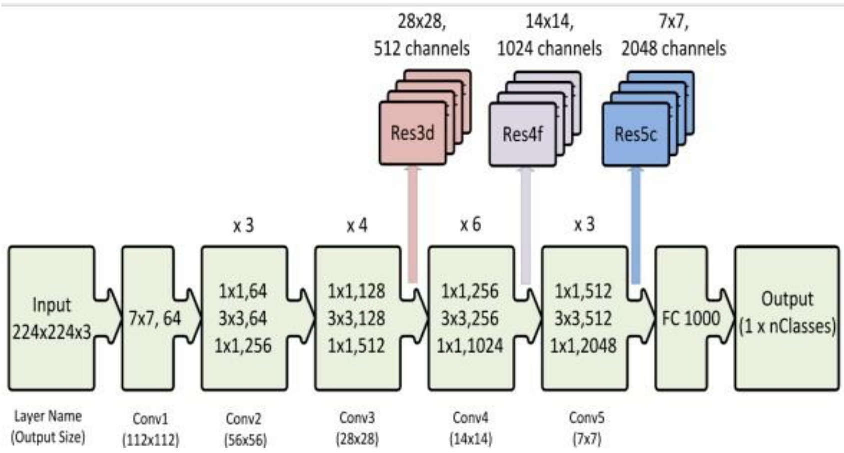

Figure 2: Resnet50 architecture

One of the CNN models used is Res Net 50. Initially, ResNet50 CNN architecture using transfer learning with pre-trained ImageNet weights is used. The standard ImageNet weights classify the objects into 1000 categories, so the top layer is excluded, and some additional layers were added in the model. The standard input image size for ResNet50 is 224 × 224 × 3. As a Classifier, the top layer is removed and then the following layers are added to the original model. Flatten Layer, Softmax Layer (with 5 classes). The original layers were freezed and only the last layer is trained. The method works as it is found that the kind of information needed to distinguish between all the 1000 classes in ImageNet is often also useful to distinguish between new kinds of images (fundoscopic retinal images in our case). In Fine tuning, the top layer is removed and then the following layers were added to the original model.

- Flatten Layer
- Fully Connected Layer (512 neurons and relu activation)
- Dropout Layer (50 percent dropout and relu activation)
- Fully Connected Layer (256 neurons and relu activation)

- SoftMax Layer (with 5 classes)

Here the original layers were freeze and only the last 5 layers that were added were trained to make the model adapt to our input. Adam optimizer and categorical cross entropy as loss function is used in both cases. Categorical cross entropy is used, and the labels are fed in one-hot encoded format. The model is trained for 20 epochs and a batch size of 32 with 8000 training and 2000 validation samples in each epoch.

InceptionV3

Here, standard InceptionV3 CNN architecture with pre-trained ImageNet weights were used. Since the standard model is required to have an Image of 299 × 299 × 3 pixels, Image resizing is done. In the architecture, Softmax Layer is added with 5 output classes and placed instead of the top layers.

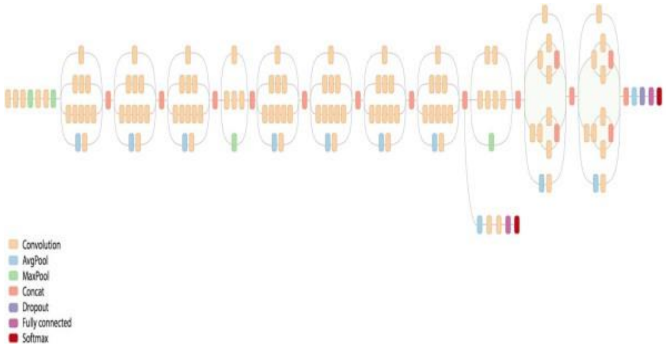

Figure 3 InceptionV3 Architecture

In ResNet50 without fine tuning no additional layers were added, the original layers were freezed and only the last layer, which is the soft max layer, were trained. The method worked as it was able to distinguish between all the 1000 classes in ImageNet is also useful to distinguish between new kinds of images and able to get an accuracy around 80%. To improve the accuracy of the ResNet50, fine tuning of the model was done where 5 additional layers were added to the original model and the additional layers were trained to make the model adapt to our input. Because of fine tuning the accuracy increased to around 84%. Also another model Inception V3 was trained to improve accuracy, here pre-trained image Net weights were used, the images were resized and the architecture was customized by adding a soft max layer with 5 classes. Hence the accuracy was improved by around 88%.

Table no 2: Results of CNN

Model	Train Accuracy	Test Accuracy
ResNet50 without Fine Tuning	80.24%	79.33%
ResNet50 with Fine tuning	84.44%	83.2%
InceptionV3 without Fine tuning	88.56%	87.17%

IV.CONCLUSION

Across all models, using CLAHE for histogram equalization improved the accuracy by a great difference (10%). This is because it made the features like microaneurysms, exudates, hemorrhage, and blood vessels more prominent. The best results were obtained using the InceptionV3 model with ImageNet pretrained weights and a SoftMax layer with 5 output units at the end, training the additional layer and the top one block of the base model. As there are very few images for classes 3 and 4 so our model got less information about these classes. Therefore, either we need more data for these classes, or we need to augment the data. Augmentation is flipping, rotating, altering brightness, zooming etc. to increase the samples artificially. Limitations of computational resources were one of the challenges and deep CNN models take a long time to run. Also, preprocessing and data augmentation techniques could have been explored. Due to time and computational constraints, there was a tradeoff between training time and accuracy.

.

REFERENCES

1. *Cao, W., Shan, J., Czarnek, N., & Li, L. (2017, November). Microaneurysm detection in fundus images using small image patches and machine learning methods. In 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 325-331). IEEE.*
2. *Shan, J., & Li, L. (2016, June). A deep learning method for microaneurysm detection in fundus images. In 2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE) (pp. 357-358). IEEE.*
3. *Goel, S., Gupta, S., Panwar, A., Kumar, S., Verma, M., Bourouis, S., & Ullah, M. A. (2021). Deep Learning Approach for Stages of Severity Classification in Diabetic Retinopathy Using Color Fundus Retinal Images. Mathematical Problems in Engineering, 2021.*
4. *Qomariah, D. U. N., Tjandrasa, H., & Fatichah, C. (2019, July). Classification of diabetic retinopathy and normal retinal images using CNN and SVM. In 2019 12th International Conference on Information & Communication Technology and System (ICTS) (pp. 152-157). IEEE.*
5. *ATILA, M., LACHGAR, M., HRIMECH, H., & KARTIT, A. (2021). Diabetic Retinopathy Classification Using ResNet50 and VGG-16 Pretrained Networks. International Journal of Computer Engineering and Data Science (IJCEDS), 1(1), 1-7.*
6. *Yaqoob, M. K., Ali, S. F., Bilal, M., Hanif, M. S., & Al-Saggaf, U. M. (2021). Resnet based deep features and random forest classifier for*

diabetic retinopathy detection. Sensors, 21(11), 3883.

7. *Carrera, E. V., González, A., & Carrera, R. (2017, August). Automated detection of diabetic retinopathy using SVM. In 2017IEEE XXIV international conference on electronics, electrical engineering and computing (INTERCON) (pp. 1-4). IEEE.*
8. *Elswah, D. K., Elnakib, A. A., & Moustafa, H. E. D. (2020, September). Automated diabetic retinopathy grading using resnet. In 2020 37th National Radio Science Conference (NRSC) (pp. 248-254).IEEE.*
9. *Bhardwaj, C., Jain, S., & Sood, M. (2021). Diabetic retinopathy severity grading employing quadrant-based Inception-V3 convolution neural network architecture. International Journal of Imaging Systems and Technology, 31(2), 592-608.*
10. *Bhardwaj, C., Jain, S., & Sood, M. (2021). Deep Learning–Based Diabetic Retinopathy Severity Grading System Employing Quadrant Ensemble Model. Journal of Digital Imaging, 34(2), 440-457.*
11. *Elsawah, D., Elnakib, A., & Moustafa, H. E. D. S. (2020). Deep Learning Grading System for Diabetic Retinopathy using FundusImages. (Dept. E). MEJ. Mansoura Engineering Journal, 45(4), 1-8.*
12. *Sopharak, A., Uyyanonvara, B., & Barman, S. (2013). Automated microaneurysm detection algorithms applied to diabetic retinopathy retinal images. Maejo International Journal of Science and Technology, 7(2), 294.*
13. *Van Grinsven, M. J., van Ginneken, B., Hoyng, C. B., Theelen, T., & Sánchez, C. I. (2016). Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection incolor fundus images. IEEE transactions on medical imaging.*
14. *Gurudath, N.; Celenk, M.; Riley, H.B. Machine learning identification of diabetic retinopathy from fundus images. In Proceedings of the 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, USA, 13 December 2014; pp.*