Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 1055–1066

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Diabetic Retinopathy Grading From Color Fundus Images: An Autotuned Deep Learning Approach

Athira T R , Jyothisha J Nair*

Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India.

Abstract

Diabetic Retinopathy (DR) is a consequence of long-term diabetes which affects the eyes. It causes affliction to the veins in the eyes causing raptures on the retina which can impact vision. If the condition is failed to be detected at an early stage, it can lead to complete vision loss. The conventional diagnosing cycle of DR using fundus images requires very skilled practitioners due to the minute nature of features of the anomalies which can even lead to misdiagnosis and is time consuming. Therefore, devising an automated method for the diagnosis of DR can assist individuals with diabetes to identify symptoms of DR at a very early stage. This research also encompasses the classification of the detected images into its corresponding 3 stages namely No Diabetic Retinopathy (No DR), Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR) that can greatly aid in monitoring the dynamics of the key features including lesions, hemorrhages, and density of blood vessels. Deep learning algorithms has gotten to become a well-known method that can accomplish a wide variety of classification tasks. However, most of these methods are only efficient in classifying the various stages of DR with low accuracies notably, for the early stages. The devised algorithm in this research uses enriched images processing techniques, automatic hyperparameter tuning and neural network training strategies to provide more emphasis on the minute features for better prediction. The algorithm was tested and compared with modified Resnet50, VGG16, Mobilenetv2, Inceptionv3 and InceptionResnetv2 which gave a classification accuracy of 94.7%, 86.1%, 85.8%, 85.3% and 87% respectively with corresponding detection accuracies of 99.8%, 94%, 94.2%, 94.9% and 98.2% respectively on a test set of 508 images. Using the proposed algorithm, the results indicate Resnet50 based network gave superior performance for both detection and classification tasks.

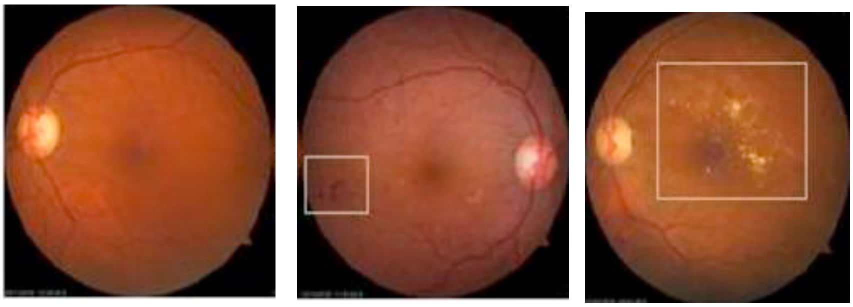
Keywords: Diabetic Retinopathy; Convolutional Neural Network; Non-Proliferative Diabetic Retinopathy; Proliferative Diabetic Retinopathy

* Corresponding author. Tel.: +91-9447205116
E-mail address: jyothishaj@am.amrita.edu

10.1016/j.procs.2023.01.085

1. Introduction

Diabetic Retinopathy is an extreme complication of diabetes mellitus which even in the current days remains as the leading cause of blindness globally [8]. Countries of huge population like China, Indonesia, India, and Bangladesh shares 45% of the global counts in diabetes [1]. As the numbers are anticipated to move up, an automatic approach for clinical diagnosis will be of much help. The primary motivation of this research is to come up with an application for state of the art deep learning algorithms in the medical image analysis and diagnosis to aid people with long term diabetes to identity and take precautionary actions against its progression. An automatic system would help people with diabetes to identify the signs of the disease at an earlier stage which can highly reduce the clinical load on retina experts and help monitor the dynamics of the lesions. Thus, the research on automated diagnosis of DR becomes more and more crucial in the past few years. Through this research we propose an efficient algorithm for detection and a 3-stage severity grading of DR. While most research do not give much emphasis on blood vessels the proposed approach gives more importance to feature enhancement techniques that enhance blood vessels treating it as an important feature. This research has also eliminated the need for manual hyperparameter tuning by incorporating autotuners to find best possible parameters from a large search space and to dynamically add extra layers during runtime. In this research a Convolutional Neural Networks approach is proposed to automate the method of DR screening using preprocessed fundus retinal image as input.



(a) (b) (c)

Fig. 1. (a) Normal image (b) Red lesion (c) Exudate

Classes in DR classification include No DR, NPDR and PDR [7]:

- No Diabetic Retinopathy: At this stage no signs of DR can be seen at this stage.
- Non-Proliferative Diabetic Retinopathy is the earliest stage in DR where micro aneurysms such as red lesions start to occur. The potential of blood transportation due to blood vessel distortion and swelling decreases as the disease progresses. NPDR is characterized by the presence of minute red lesion in a scale which is hard to be identified on visual inspection as depicted in Fig1b making it difficult to detect the disease at this stage. Vision is not usually affected at this stage although there are higher risks for developing eyesight issues in the future. If the disease is identified at this stage oral treatment can help reverse the effect.
- Proliferative Diabetic Retinopathy is a highly progressed stage of DR, when the retinal blood vessels cause activates proliferation of the newly formed blood vessels emerging inside in the vitreous gel in the eye. Exudates becomes more prevalent at this stage. Laser treatment may have to be prescribed to avoid complete vision loss. The slight yellow patches depicted inside the localized region in Fig1c shows an example of exudates.

Each stage has its characteristics and numerous properties which in certain cases are hard for doctors to possibly take note of, thereby increasing the chances of incorrect diagnosis. This led to the need for devising an automated approach for DR detection.

Object detection and classification utilizing different machine learning procedures have been a vital focal point of the exploration for the research society [2]-[3]. Particularly with the approach of CNN, different models have been proposed to achieve the errands of computer vision (CV), speech recognition, natural language processing (NLP and advanced mechanics, [4]. There are also different instances of deep learning (DL) use in biomedical applications [5],[18]-[23],[27],[28]. In this research, we have presented a strategy that covers the data preparation, recognition, and classification of DR from fundus images. The primary focus by incorporating deep learning techniques in image classification is to introduce efficient algorithms to replace the dependence hand coded key points for classifying DR images.

Acharya et al. proposed a novel automated technique for identifying the DR stages using simple image processing techniques. The algorithm gave an average accuracy, specificity and sensitivity as 85%, 86% and 82% respectively. This result could be increased by the usage of more efficient feature enhancement techniques [9]. Sudeshna et al. has done considerable work in extracting out blood vessel and removing optic disc, which in most cases may result in creation of residues leading to a false detection. The research discussed techniques of morphological closing and opening for the extracting out natural structures in the retina. But the similarity in the nature of color characteristics for normal and abnormal features of in the retina can sometimes result in residues [10].

Feature extraction is one important step in any kind of disease detection. With the use of CNNs manual feature extraction becomes no longer necessary, a single CNN network can automatically extract features and do classification. The findings of several other works related to this research is consolidated in table 1.

Table 1. Summary of related works.

Sl No	Authors Name	Method	Observations
1	Aacharya et.al, 2008 [9]	Proposed automated technique to identify DR stages using simple image-processing and data-mining techniques using higher order spectra.	Simple image processing gives moderate accuracy.
2	Doshi et.al, 2016 [16]	GPU accelerated deep CNN based algorithm.	Image processing techniques can be used for decreasing overfitting through increasing complexity of the data.
3	Wang et.al, 2018 [13]	Three convolutional neural networks were separately analyzed.	Paper shows InceptionNet achieved the highest accuracy.
4	Sudeshna et.al, 2017 [10]	Extracted and removed all features that are possible candidates for false negatives.	The approach resulted in enhanced the performance compared to existing approaches but performs slightly weaker in detecting red lesions.
5	Wen et.al, 2020 [14]	Designed a novel Resnet-50 based TCNN for fault diagnosis.	Explored the capabilities of Resnet based model.
6	Alban et.al, 2016 [15]	Implemented DR classification using Google Net, AlexNet and Baseline (Custom Architecture).	The model performs well in comparison to human evaluation metrics and the performance of weak learners can be boosted with the approach of ensemble learning.

The rest of the paper is organized into various sections which mainly focus on methodology, results discussion, conclusion and future work. The major steps in the implementation of the proposed algorithm starting from details of the dataset to the training strategies are included in section 2. Section 3 covers the experimental setup and result evaluations. Section 4 is a brief discussion of the paper. Finally, section 7 summarizes the conclusions and directions for future enhancements.

2. **Methodology**

The proposed algorithm in this work utilizes enriched preprocessing and highly optimized parameter tuning for modelling and training to ensure the best possible results. The whole setup can be broken down into preprocessing, hyperparameter optimization, hybrid network training and finally analysis of the classification results from various networks. The basic blocks of the proposed algorithm is summarized in figure 2. The dataset contains the images of each of the 3 stages in DR. Image resizing, grey scaling, gaussian blurring and circular cropping were the image preprocessing steps used to enhance the important features in the images. Preprocessing was followed by hyperparameter tuning for each selected base network which included training parameters as well as model building parameter like addon layers. Instead of using a conventional manual tuning approach for hyperparameter optimization, we implemented an hyperparameter optimizer using SLSQP to automize the whole process as well as to make sure to always get the best set of hyperparameters from a search space. The best set of parameters for each of the 5 models were used for the final model building and training. The classification results of each of the 5 models are then analyzed and compared. The whole research was done on a Nvidia Tesla P100 GPU with a maximum RAM usage of 24 GB as high RAM usage was necessary during preprocessing and auto hyperparameter tuning to avoid crashes and to increase the speed of execution.

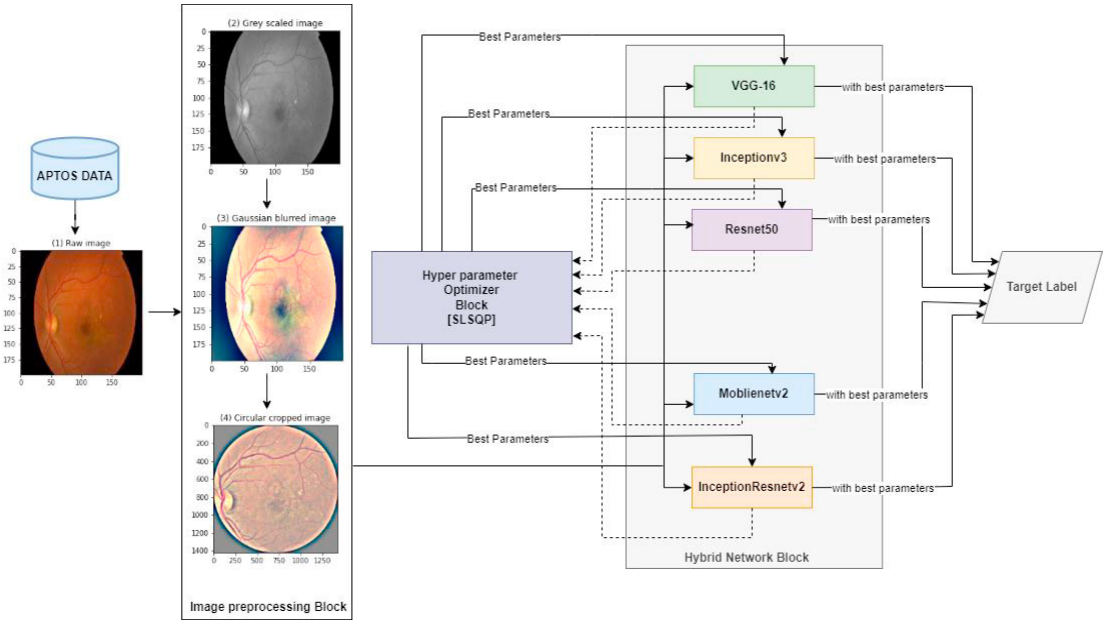


Fig. 2. Schematic Diagram of the Proposed Approach

2.1. Dataset

The dataset used is a subset of the Asia Pacific Tele Ophthalmology Society. The complete dataset consists of 18590 fundus images. The dataset is relatively inhomogeneous in terms of quality as these images were not obtained in a controlled lab setup and used different cameras to capture resulting in different image resolutions [14]. To overcome these challenges, the images in APTOS dataset were screened additionally by ophthalmologist whenever required for confirmation. Therefore, subset of 3386 images were used for the train and test datasets and is reported in this paper. 85% of the dataset was utilized for training, the rest 15% was utilized for testing. The classes in the data includes No DR (class 0), NPDR (class 1) and PDR (class 2).

2.2. Preprocessing

Preprocessing techniques are chosen to provide more focus in enhancing the features like blood vessel, red lesions and exudates. Due to the huge amount of data summing close to 10 GB, to reduce the time for preprocessing thread pooling was used. Thus, the whole process of preprocessing was divided into 4 different threads and ran on 4 different cores concurrently.

2.2.1. Resizing the images

As the data was collected from various sources at different time period, each image data had differing image size and storage memory. Therefore, rescaling all the images in the entire dataset to a uniform image size was essential for the better computation capability. Figure 1 shows the visualization of 3 random images, chosen from the three different classes. It was observed that the features such as red lesions and exudates were not clearly visible, especially in the images which lacked proper lighting.

2.2.2. Convert to grey scale

Creating a better model, requires training the model on a uniform well defined dataset. The first technique used for feature enhancement was converting the color images to grey scale images. The features such as exudates (yellow spots) was portrayed very well, whereas the feature such as red lesions were barely visible. Training the model on such dataset will result in very high bias and high training error. A far better result was observed using advanced feature enhancement techniques such as gaussian blur and circular cropping.

2.2.3. Gaussian Blur

It is an image processing technique used to lessen the amount of noise in an image. An image, which is a huge two-dimensional matrix of numbers representing the color of the pixel is passed through the gaussian filter. It is usually accomplished by convolving pixel array with the gaussian filter. A 2-dimensional gaussian filter can generally be expressed as:

$$G_{2D}(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (1)$$

Here x and y are the location indices and standard deviation of the distribution being denoted by σ. The idea of σ controls the fluctuation around a mean of the gaussian distribution, which decides the degree of the obscuring impact around a pixel [11]. A wide range of combinations of these parameters were experimented to identify the parameter set with the best result. The gaussian blur utilized that gave the best blood vessel and red lesion enhancement result had a sigma x which is the gaussian kernel standard deviation in x direction as 30. After applying gaussian filter, the areas that contained more information were extracted. To sharpen the edges of the image obtained after gaussian blur. A weighted masking is used to sharpen the edges of the image obtained after gaussian blur. The first image passed is the resized image with alpha weighted as 5 and the second image passed is the image after applying gaussian blur with beta weighted as -5. The gamma value is specified as 128. Images with both features of exudates and red lesions and more importantly blood vessel were enhanced more than the original images were obtained as shown in the third step of figure 3.

2.2.4. Circular cropping

To give more focus to the part containing Blur the fundus image of the eye, a circular cropping is supplied to all the images to discard nonessential background hence making a uniform training dataset. This helps to fit the images

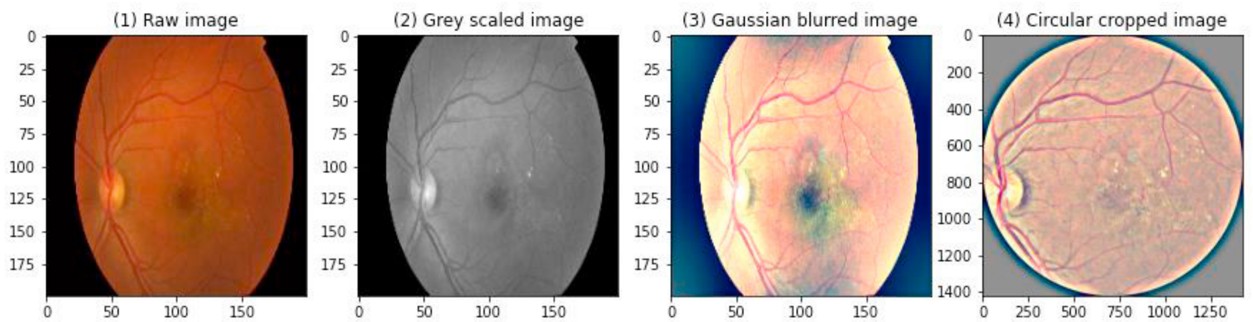


Fig. 3. Outputs of various stages in preprocessing

to the new image frame secured after removing the background through the circular cropping technique as shown in the figure 3. In order to apply circular cropping on a color image, the image is first converted to grey scale and a mask is created which would pass only those pixel values of the image, which are greater than the specified tolerance value. In cases where the image is too dark, we return the original image to avoid the cropping of the entire image. In the other case, the boolean array obtained after applying the masking can be used to index the image for extracting required bounding boxes making use of broadcasted indexing and a bitwise and operation is done between this image and the dummy circle created for doing the circular cropping.

2.3. Visualizing Features

Feature space visualization was done to identify the degree of model complexity to choose the right model for the classification task. A non-linear technique of feature visualization called t-Distributed Stochastic Neighbor Embedding (t-SNE) was for information investigation and imagining high-dimensional information in an unsupervised manner in this study. Unlike PCA this technique is more than just a mathematical procedure but is a probabilistic technique. t-SNE limits the difference between two dispersions: one that that estimates pairwise likenesses of the input objects and the other that estimates the pairwise likenesses of the low-dimensional data in the embedding. [12]. Plotting t-SNE gave us an idea on the level of difficulty in separating the various classes in data. From the figure 4 it can be noted that the classes within DR specifically, classes 0 and 2 have many overlapping features hence difficult to distinguish. t-SNE plots proves that even though there are principal features that are comparatively easy separable in class 0 there are a group of other features in class 0, 1 and 2 which are highly convoluted and hard to separate out. The main feature of t-SNE is that it tries to deconvolute relationships between neighbors in high-dimensional data giving us a better understanding of the separability of the classes and is a reducing the dimensions for visualization.

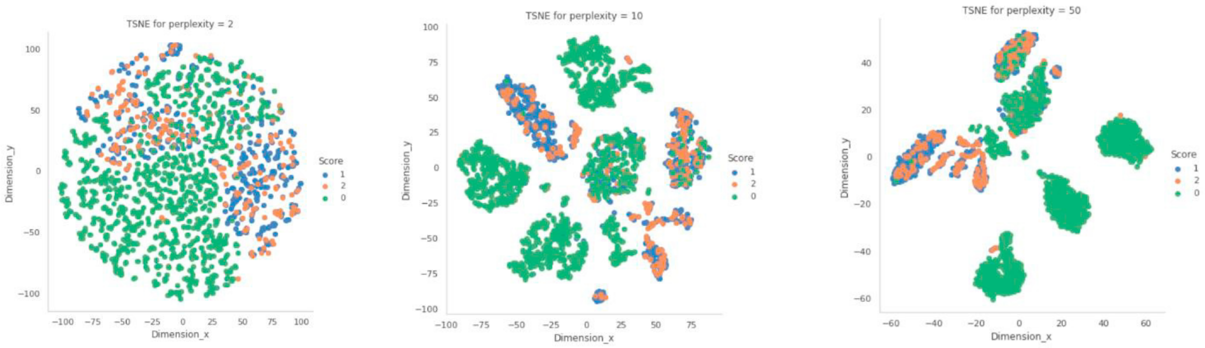


Fig. 4. 2-Dimensional visualization of data with t-SNE

2.4. Training

Modified Resnet50, InceptionResnetv2, Mobilenetv2, Inceptionv3 and VGG16 were used for training. The technique of transfer learning was employed to get the maximum out of it. After the elaborate training processes, it was observed that Resnet based networks viz. Resnet50 and InceptionResnetv2 gave better performances compared to the non-Resnet based networks. Resnet50 modified network gave the maximum accuracy among all the tested networks. ResNet50 is a convolutional neural network with 50 layers. The pretrained variant of the network on ImageNet data is used in the in this study [1]. Resnet is characterized by the presence of skip connections that greatly reduces the effect of vanishing and exploding gradients. Resnet50, Mobilenetv2 and InceptionResnetv2 are the other networks that uses skip connections in the basic blocks a sample of which is depicted in figure 5.

To the base Resnet model an additional block of 4 extra layers were added to get the final model for training. A dropout of 0.3 was used for transfer learning. The final dense activation layer was of sigmoid with 3 outputs for the 3 different classes. The performance of the deep learning models is extremely sensitive to altering hyperparameters hence we created an hyperparameter optimizer module to test and tune all possible combinations of hyperparameters from its possible sample space. Hence the best hyperparameters were found using auto tuners with a well-defined search space, The auto-tuner used for the experimentation was a ray tuner. A tuner supports numerous search algorithms viz PBT, Random Search, ASHA, Bayesian Optimization, Hyper Band, Median Stopping, BOHB. The best hyperparameters were chosen for training.

Table 1. Addon layers in each model

Sl No	Resnet 50	Inception Resnet v2	Mobilenet v2
1	Dense + activation = 'relu'	Global Average Pooling2D	Dense Layer
2	Dropout [ratio - 0.3]	Dense + activation = 'relu'	activation = 'sigmoid'
3	Dense + activation = 'relu'	Dense + activation = 'sigmoid'	
4	Dense + activation = 'sigmoid'		.

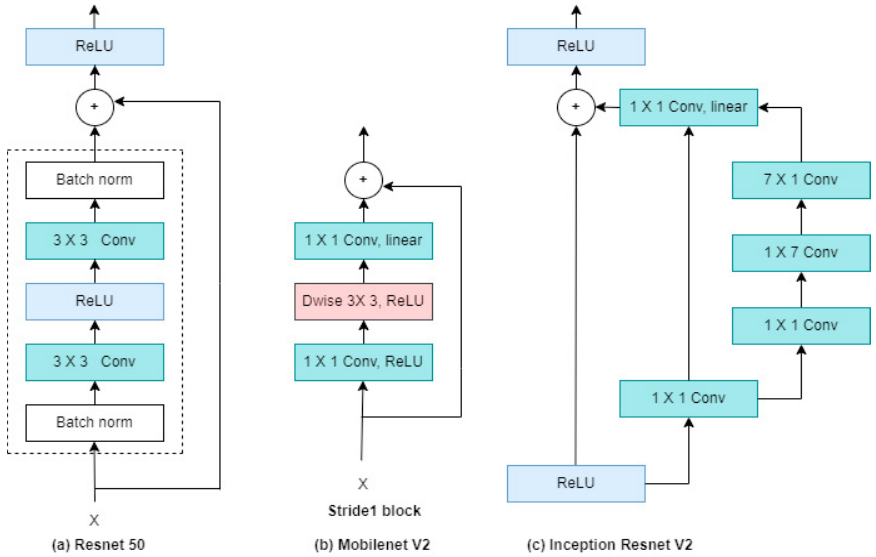


Fig. 5. Basic building blocks of chosen base networks with skip connections

The automatic hyperparameter tuner picks up the best set of hyperparameters from a well-defined search space thereby making sure to get the best combination of hyperparameters. The parameters that were allowed to be chosen by the tuner include batch size, learning rate, number of addon layers, inclusion or exclusion of batch normalization layers, number of epochs to train the choice of enforcing positive bias and weight constraints.

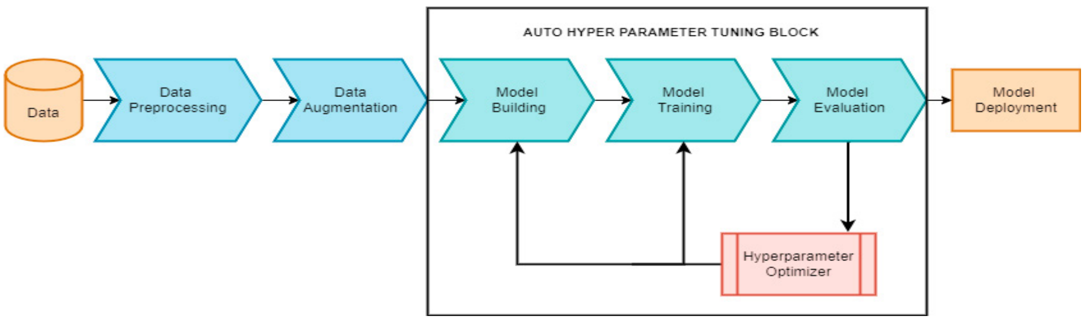


Fig. 6. Hyperparameter Tuning Process

Figure. 6 shows the autotuner block in the workflow. Figure 7 shows the basic algorithm used to implement the hyperparameter tuner [17][24].

Algorithm 1: Basic algorithm of hyperparameter tuning

Input *(Hyperparameter space θ, Search algorithm S, Evaluation function E (θ), max number of evaluations* n_{max}*)*
Select initial set of hyperparameters $\theta_0 \in \theta$
Evaluate initial score $yo = E(\theta_0)$
Set $\theta^* = \theta_0$ *and* $y^* = E(\theta_0)$
For $n=2\ldots n_{max}$
Select new set of hyperparameter $\theta_{new\,from}$ *defined space using defined Search algorithm S*
Calculate E_{new} *with new parameters to get* $y_{new} = E(\theta_{new})$
If y_{new} *< Threshold error (extremely small value)*
Assign $\theta^* = \theta_{new}$ *and* $y^* = y_{new}$
Return $\theta^* = \theta_{new}$ *and* $y^* = y_{new}$

Fig. 7. Algorithm of hyperparameter tuning

It is prescribed to utilize batch sizes with powers of 2 since it fits with the memory of the CPU. For transfer learning, an extremely low learning rate is favored with the goal that it doesn't change a lot of what is recently learnt by the network. The quantity of addon layers relies on the amount of learnt patterns that can be transferred from the layers of the pre-trained model. Involving every one of the layers for transfer learning, a basic flatten layer and dense layer with softmax is sufficient however since we integrated the feature extraction it required more layers toward the end. A few Optimizers were experimented including stochastic gradient descent (SGD) and root mean squared propagation (RMSprop). SGD with an exceptionally low learning rate expected more iterations to finish training the model with a reasonable amount of pattern being learnt. Thus, RMSprop was utilized to come by the expected outcome. Batch normalization and dropout layers are placed in between the dense layers to avoid any chances of overfitting. Training was done in 4 steps with 50 epochs in each step. The initial 2 epochs were trained on a very low learning rate as a warmup. The whole network was trained 200 epochs excluding the epochs for warmup with a slightly higher learning rate of 0.001 finally to yield a training accuracy of 98.9%.

Table 2. Autotuner model building and training parameter space

Sl No	Hyper parameter	Optimizer search space
1	Epochs	[50, 100, 150, 200, 500, 1000]
2	Batch Size	[4, 8, 16, 32, 64]
3	Learning Rate	[10−1, 10−2, 10−3, 10−4, 10−5, 10−6]
4	Warmup Epochs	[1, 2, 5, 10, 15]
5	Image Height, Width	[126, 224, 512]
6	Early stopping Patience	[5, 8, 10, 12, 15]
7	RLprop Patience	[2, 3, 4, 5, 8, 10]
8	Decay Drop	[0.1, 0.2, 0.3, 0.4, 0.5]
9	Addon Dense layers	[1, 2, 3, 4, 5, 6]
10	Addon Dropout layers	[1, 2, 3, 4, 5, 6]
11	Dropout ratio	[0, 0.1, 0.2, 0.3, 0.4, 0.5]

There are four main approaches used during training - A warmup phase, training with higher learning rate, reducing learning rate on plateau and early stopping.

2.4.1. Warmup phase

Warm up step is used for reducing the learning rate to decrease the effect of deviating the model from learning the immediate new data set to which it is exposed. It helps lower the impact of the primacy effect due to the early training examples. In the absence of a warmup phase, it might be required run extra epochs to get the desired convergence. The first 2 epochs were allocated for warmup training with a learning rate of 0.0001. It also helps fight early overfitting [6].

2.4.2. Training with higher training rate

After training 2 epochs with lower learning rates the rest of the 200 epochs are trained with higher learning rates of 0.001 after the warmup phase for the Resnet50 based network. Hyperparameter tuners were used to identify the optimal parameters from a large search space to ensure the best accuracy.

2.4.3. Reducing Learning rate on plateau

This approach was for reducing the learning rate when the matric stops improving. We monitored the validation loss of consecutive 3 epochs for creating a decay drop in case of not much improvement in the metric measured.

2.4.4. Early stopping

Early stopping is an effective strategy to avoid overtraining. The training process is halted as soon as the performance on the current validation cycle diminishes comparing the last validation data performance. Using the above stated strategies, Resnet50 based model was able to attain a training accuracy of 98.9%. with a validation accuracy of 90.88%.

3. Results

The experimentation was done in two different approaches to separately evaluate all the 5 models for their disease detection and disease classification abilities. 508 images were chosen as the test set for both the testing. Initially a 3-class testing with classes No DR, NPDR and PDR was done to assess the proficiency of the model on grading the severity of the disease. A second level of testing was done to evaluate the disease detection capability of the models. To test the detection capability all class 0 images were as taken as the non-DR class and data from class 1 and 2 as DR class. The data was tested on modified Resnet50, InceptionResnetv2, Mobilenetv2, Inceptionv3 and VGG16.

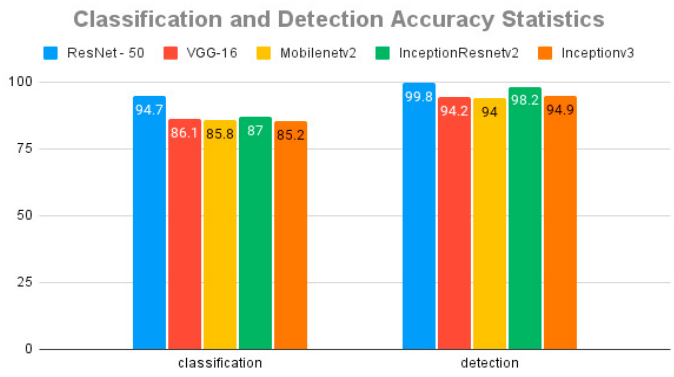


Fig. 8. Analysis of overall accuracy on experimented networks

A total of 4 evaluation matrices were used to assessing the model's performance on unseen data which includes, accuracy, precision, recall, f1 score and kappa score. These results are consolidated in table 3. As this is a multi-class problem the overall values of the evaluation matrices are found in three ways as micro, macro and weighted for the best network is consolidated in table 4. The disease detection testing showed an accuracy of 99.8% and disease classification 94.7% on Resnet50 based network as shown in figure 8.

Table 3. Overall Comparison table for detection and classification based on accuracy and kappa score

Evaluation Matric	Resnet 50	VGG − 16	MobilenetV2	InceptionResnetV2	Inception V3
Classification Accuracy	0.947	0.861	0.858	0.870	0.852
Kappa Score	0.884	0.713	0.739	0.929	0.724
Detection Accuracy	0.998	0.942	0.94	0.982	0.949

Table 4. Detailed analysis of Resnet 50 based network

	Detection			Classification		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Micro	0.998	0.998	0.998	0.94	0.94	0.94
Macro	0.999	0.997	0.998	0.92	0.81	0.83
Weighted	0.998	0.998	0.998	0.94	0.94	0.93

The analysis of results with the same algorithm done for 4 other major networks is consolidated in figure 8 and figure 9. Using the above algorithm Resnet50 based network is tested to show an overall higher performance in terms of all the evaluated parameters taking all the 3 classes together. It can also be noted that InceptionResnetv2 showed higher performance considering precision of class 2 alone thereby opening possibilities of even higher performance by combining the 2 networks. This result can be attributed to the architecture of Resnet based networks. The presence of skip connections helps propagate the importance of minute features without being degraded by activation functions or other mathematical computations.

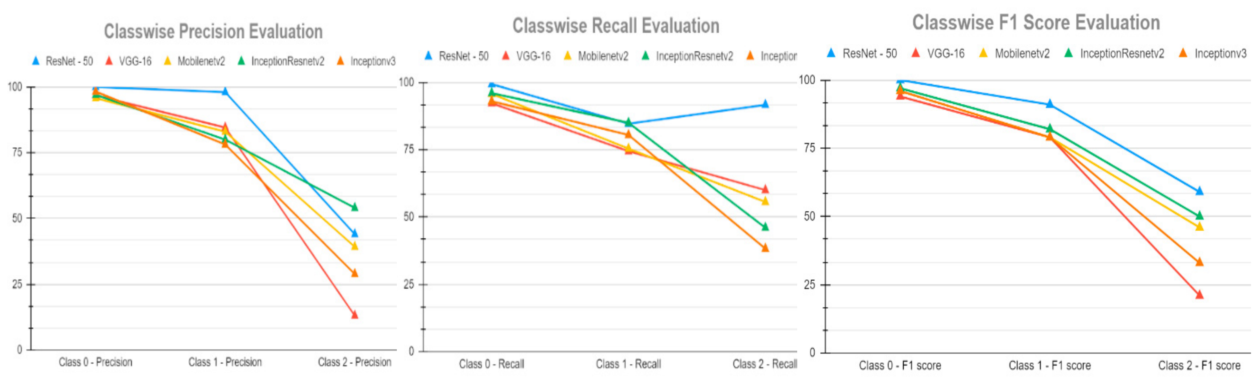


Fig. 9. Classwise Analysis of performance parameters of experimented with major networks

Rao at al. experimented with 5 different state-of-the-art deep learning models with Adam and Stochastic Gradient Descent (SGD) on 2 different input sizes viz 224x224 and 512x512 to give promising results as depicted in Table 5 and 6. On comparison with the work done by Rao at el. For the networks common in both the studies for detection both gave higher performance with the Resnet50 network. With our approach Resnet50 and inception v3 gave significantly higher performance and VGG16 gave extremely closer performance to that of Rao at el. Analysis of the classification results of both the research shows that Rao at el. achieved a maximum accuracy of 88.1% on InceptionResnetv2, and our work achieved 94.7% with the much lighter Resnet50 based network.

Table 5. Detection performance comparison of proposed approach with prior works

	Proposed			Rao et al [26]		
	Resnet 50	VGG 16	Inceptionv3	Resnet 50	VGG 16	Inceptionv3
Accuracy	0.998	0.942	0.949	0.966	0.95	0.896
Precision	0.99	0.94	0.95	0.97	0.95	0.89
Recall	0.99	0.95	0.95	0.97	0.95	0.89
F1 Score	0.99	0.95	0.95	0.965	0.950	0.892

Table 6. Classification performance comparison of proposed approach with prior works

	Proposed				Rao et al [26]			
	Resnet50	VGG16	Inceptionv3	Inception Resnetv2	Resnet50	VGG16	Inceptionv3	Inception Resnetv2
Accuracy	0.947	0.861	0.852	0.87	0.785	0.747	0.812	0.881
Precision	0.94	0.83	0.84	0.86	0.79	0.76	0.83	0.88
Recall	0.94	0.85	0.85	0.87	0.78	0.76	0.81	0.88
F1 Score	0.93	0.83	0.84	0.87	0.778	0.756	0.805	0.882

4. Discussion

This research proved that machine learning algorithms such as neural networks do have high future scope in detecting abnormalities from medical images. Research have proactively demonstrated the proficiency of CNN of CNN technique in tasks like object detection. With the results from this research, it is evident that the scope of neural networks is not just confined in finding object of considerable sizes and definite shapes, but neural networks can also be successfully used for classifying objects with very tiny features with irregular shapes and sizes. The test accuracy of 99.8% for detection and 94.7% on classification obtained using Resnet based networks proves that with the right kind of image enhancement in preprocessing CNN can be found to be highly proficient for tasks like abnormality detection and classification from medical images.

5. Conclusion and Future Work

From the research it can be observed that networks with skip connections performed well in classifying the stages when the feature are extremely small. Resnet50 based network topped the overall performance matrices for both detection and classification proving that with the right set of hyperparameters and preprocessing light weight skip connected networks can be efficiently utilized for DR detection and classification. The proposed system can greatly reduce the time of diagnosis and thus improve the quality of life of diabetic patients. Improvements in this works can be made in a several viewpoints. The 3-level classification can be extended to a more elaborate 5 class problem making use of more advanced features. An ensemble approach can be implemented to boost the overall proficiency of the algorithm. The power of CNN can be leveraged for optic disc removal and blood vessel extraction for enhancing classification performance without causing residues. CNN based image semantic segmentation for cotton wool natured exudates which are hard to detect using normal image processing methods could likewise be examined to carry out fine-grained DR classification.

References

[1] D. S. Fong, F. L. Ferris III, M. D. Davis, E. Y. Chew, E. T. D. R. S. R. Group. (1999) "Causes of severe visual loss in the early treatment diabetic retinopathy study: Etdrs report no. 24", *American journal of ophthalmology* **127 (2)**: 137–14

[2] M. Jones, P. Viola. (2003) "Fast multi-view face detection.", *Mitsubishi Electric Research Lab TR-20003-96* **3 (14)**: 2.

[3] W. Cheung, G. Hamarneh. (2009) "n-sift n-dimensional scale invariant feature transform", *IEEE Transactions on Image Processing* **18 (9)**: 2012–2021.

[4] A. Krizhevsky, I. Sutskever, G. E. Hinton. (2012) "Imagenet classification with deep convolutional neural networks", *Advances in neural information processing systems* **25**.

[5] L. Yang, Y. Zhang, J. Chen, S. Zhang, D. Z. Chen. (2017) "Suggestive annotation: A deep active learning framework for biomedical image segmentation", *in: International conference on medical image computing and computer assisted intervention*, Springer: 399–407

[6] F. He, T. Liu, D. Tao. (2019) "Control batch size and learning rate to generalize well: Theoretical and empirical evidence", *Advances in Neural Information Processing Systems* **32**.

[7] W. R. Memon, B. Lal, A. A. Sahto. (2017) "Diabetic retinopathy; Frequency at level of hba1c greater than 6.5%", *The Professional Medical Journal* **24 (02)**: 234–238.

[8] A. W. Stitt, T. M. Curtis, M. Chen, R. J. Medina, G. J. McKay, A. Jenkins, T. A. Gardiner, T. J. Lyons, H.-P. Hammes, R. Simo. (2016) "The progress in understanding and treatment of diabetic retinopathy", *Progress in retinal and eye research* **51**: 156–186.

[9] R. Acharya U, C. K. Chua, E. Ng, W. Yu, C. Chee. (2008) "Application of higher order spectra for the identification of diabetes retinopathy stages", *Journal of medical systems* **32 (6)**: 481–488.

[10] S. S. Kar, S. P. Maity. (2017) "Automatic detection of retinal lesions for screening of diabetic retinopathy", *IEEE Transactions on Biomedical Engineering* **65 (3)**: 608–618.

[11] S. Misra, Y. Wu. 2020 "Chapter 10 - machine learning assisted segmentation of scanning electron microscopy images of organic-rich shales with feature extraction and feature ranking", in: S. Misra, H. Li, J. He (Eds.), *Machine Learning for Subsurface Characterization,* Gulf Professional Publishing: 289–314.

[12] L. Van der Maaten, G. Hinton. (2008) "Visualizing data using t-sne", *Journal of machine learning research* **9 (11)**: 2579-2605

[13] X. Wang, Y. Lu, Y. Wang, W.B. Chen. (2018) "Diabetic retinopathy stage classification using convolutional neural networks", in: *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, IEEE: 465–471.

[14] L. Wen, X. Li, L. Gao. (2020) "A transfer convolutional neural network for fault diagnosis based on resnet-50", *Neural Computing and Applications* **32 (10)**: 6111–6124.

[15] M. Alban, T. Gilligan. (2016) "Automated detection of diabetic retinopathy using fluorescein angiography photographs", *Report of standford education*.

[16] D. Doshi, A. Shenoy, D. Sidhpura, P. Gharpure. (2016) "Diabetic retinopathy detection using deep convolutional neural networks", *in: 2016 International Conference on Computing, Analytics and Security Trends (CAST)* IEEE: 261–266.

[17] B. G. Galuzzi, I. Giordani, A. Candelieri, R. Perego, F. Archetti, (2020) "Hyperparameter optimization for recommender systems through bayesian optimization", *Computational Management Science* **17 (4)**: 495–515.

[18] R. Priya, V. Sreelekshmi, J. J. Nair, G. Gopakumar. (2021) "Breast mass classification using classic neural network architecture and support vector machine", *in: Advances in Computing and Network Communications*, Springer: 435–448.

[19] R. Dhanya, I. R. Paul, S. S. Akula, M. Sivakumar, J. J. Nair. (2020) "F-test feature selection in stacking ensemble model for breast cancer prediction", *Procedia Computer Science* **171**: 1561–1570.

[20] R. Dhanya, I. R. Paul, S. S. Akula, M. Sivakumar, J. J. Nair. (2019) "A comparative study for breast cancer prediction using machine learning and feature selection", in: *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, IEEE: 1049–1055.

[21] A. Sreekumar, K. Nair, S. Sudheer, H. G. Nayar, J. J. Nair. (2020) "Malignant lung nodule detection using deep learning", *2020 International Conference on Communication and Signal Processing (ICCSP)*: 0209–0212.

[22] L. S. Nair, R. Prabhu, G. Sugathan, K. V. Gireesh, A. S. Nair. (2021) "Mitotic nuclei detection in breast histopathology images using yolov4", in: *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE: 1–5.

[23] K. Vaishnavi, M. A. Ramadas, N. Chanalya, A. Manoj, J. J. Nair. (2021), "Deep learning approaches for detection of covid-19 using chest x-ray images", in: *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT),* IEEE: 1–8.

[24] J. Bergstra, D. Yamins, D. Cox. (2013) "Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures", in: *International conference on machine learning*, PMLR: 115–123.

[25] Diabetic retinopathy detection: identify signs of diabetic retinopathy in eye images, https://www.kaggle.com/c/diabetic-retinopathydetection.

[26] Rao, Mihir, Michelle Zhu, and Tianyang Wang. (2020) "Conversion and implementation of state-of-the-art deep learning algorithms for the classification of diabetic retinopathy", arXiv *preprint* arXiv:2010.11692

[27] Singh, Vijendra, Vijayan K. Asari, and Rajkumar Rajasekaran. (2022) "A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease", *Diagnostics* **12 (1)**: 116.

[28] Rastogi, Priyanka, Vijendra Singh, and Monika Yadav. (2018) "Deep learning and big data technologies in medical image analysis", *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*. IEEE: 60-63.