An Efficient Segmentation Based Classification of Diabetic Retinopathy Identification using CLAHE with ResNet Model

R. Rajkumar, P. Dhanalakshmi, R. Thiruvengatanadhan

Abstract*: *Diabetic retinopathy (DR) is a widespread problem for diabetic patient and it has been a main reason for blindness in the active population. Several difficulties faced by diabetic patients because of DR can be eliminated by properly maintaining the blood glucose and by timely treatment. As the DR comes with different stages and varying difficulties, it is hard to DR and also it is time consuming. In this paper, we develop an automated segmentation based classification model for DR. Initially, the Contrast limited adaptive histogram equalization (CLAHE) is used for segmenting the images. Later, residual network (ResNet) is employed for classifying the images into different grades of DR. For experimental analysis, the dataset is derived from Kaggle website which is open source platform that attempts to build DR detection model. The highest classifier performance is attained by the presented model with the maximum accuracy of 83.78, sensitivity of 67.20 and specificity of 89.36 over compared models.

Index Terms*: *Classification, DR, Segmentation, Deep Learning, Histogram

I. INTRODUCTION

Diabetic retinopathy (DR) generally occurs to patients who acquires diabetes for long time and because of retinal damage, it causes blindness. By employing the technique of fundus imaging, the DR affected retinal structure of eyes might be detected. By focusing the eye, the fundus images will be generally captured through fundus camera. The internal surface of eye is demonstrated through fundus images which comprise of fovea, retina, blood vessels, optic disc and macula. A normal retina comprises blood vessels which carries nutrients and blood needed for eye. By nature, the blood vessels are delicate and because of additional blood pressure, they may burst in diabetic patients. Through additional small blood vessels count, the diabetic retinopathy progress because of additional pressure might be found from retinal surface. For classification of various diabetic retinopathy stages like non-proliferative DR(NPDR) and proliferative diabetic retinopathy (PDR) over actual retina, the blood vessels growth might be employed as bio-marker [1].

Numerous authors had been builtmodels for effective vessel segmentation in the past decade and the retinal images are classified depending on type and severity of disease [2]. For earlier DR recognition, an automated retinopathy classification scheme is projected [3] depending on artificial neural network (ANN). To differentiate the exudates and blood vessels, morphological operators are used by this method. In addition, an accuracy of96% is attained by the methods like genetic algorithm (GA) and fuzzy c means (FCM) [4]. The technique of multilayered thresholding is projected through [5] for blood vessels segmentation in retinopathy images. For retinal structure analysis, ridgelets [6], curvelet [7] and wavelet [8] transforms are employed additionally with fundus images. For detection of hard exudates in DR images, the approach of fuzzy logic offers 99.9% as sensitivity rate. By employing multi-scale line detector, an examination of retinal vascular feature is built [9]. Through merging the model of Gaussian mixer with nearest neighborhood technique, [10] built a project called Diabetic retinopathy analysis using machine learning (DREAM) and the classification is done by employing singular vector machine (SVM). To compute infinite perimeter regularization, L2 Lebesgue integral technique is employed by [11]. For retinal image preprocessing, the method of global thresholding is used [12]. For DRIVE and STARE datasets, the vessel detection by morphological component analysis (MCA) had achieved 0.9523 and 0.959 as accuracy correspondingly [13]. Through retina tortuosity level, the premature retinopathy can be detected in blood vessels. For vessel grading, curvature-based algorithm is projected based on tortuosity levels. By employing the model of Gaussian mixture, the optic disc boundary can be computed.

The approach of deep neural network is employed in [14] which trained over massive instances of STARE, CHEST and DRIVE datasets. But, this technique needs considerable sample amount. For computer-aided screening, telemedicine system is built through employing red lesions of retinopathic images. For DR referral [15], a direct technique might be projected by classifier training. The lesions omission reduced efforts to lesion detection needs massive data pool for classifier training. The entire conventional techniques like morphological gradients, wavelets, neural networks, and other processes need to be implemented for accurate and efficient DR detection at prior probable phases. But, these approaches are computationally complex and need more classifier training.

Revised Manuscript Received on July 05, 2019.

R. Rajkumar, Assitant Professor, Department of CSE, Annamalai University, Annamalai nagar, Tamilnadu, India.

P. Dhanalakshmi, Professor, Department of CSE, Annamalai University, Annamalai nagar, Tamilnadu, India.

R. Thiruvengatanadhan, Assitant Professor, Department of CSE, Annamalai University, Annamalai nagar, Tamilnadu, India.

Automated DR detection contains several advantages like DR might be detected at prior stages effectively. The techniques of deep learning had revolutionized the domain of amazing computer vision recently. Numerous researchers had attracted towards the image classification through leveraging convolution neural networks (CNN). Studies in this domain involve feature segmentation and blood vessels [16]. For actual image classification solution, the structures of deep CNNs were projected usually and most recent studies had progressed through classification by DR fundus images. In order to solve the problem of segmentation in blood vessel, [17] used a CNN model to derive features of image. But still, few drawbacks are present in the previously developed methods. Primarily, the accuracy of the method cannot be ensured as the dataset features are derived empirically and manually. Next, the datasets are of low quality and comprise small size some fundus images with single collection environment relatively offer complexities to compare the algorithm's performance for experimental purposes. For enhancement in performance, [18] projected the framework of AlexNet. There are some superior frameworks of CNNs like GoogleNet [19] and VggNet [20]. The recently presented Residual Network (ResNet) is a significant network models that improves the CNNs performance during classification of image. To accelerate the learning time and to compare with existing methods like VggNet, AlexNet and GoogleNet, we use transfer learning that offers accurate and automatic detection with visual damages that can be reduced towards minimum degree. When comparing with existing techniques, the projected method comprises subsequent enhancements on convergence time for dataset of massive size and exhibits superior performance over classification.

Keeping the limitations of the existing models, in this study, we present an efficient segmentation based classification model. The presented model involves two main processes namely image segmentation and image classification. For segmentation purposes, Contrast limited adaptive histogram equalization (CLAHE) is used and ResNet is used for classification purposes. To validate the presented model on the DR classification process, a benchmark dataset from Kaggle website is employed. And, a set of measures namely accuracy, sensitivity and specificity are used to analyze the experimental results in a clear manner. Finally, a detailed comparative study with the recently presented models also takes place to ensure the betterment of the presented model.

II. PROPOSED METHOD

The overall process of the presented segmentation based classification model is shown in Figure 1. As shown in figure, the applied input image will undergo preprocessing stage. Then, the preprocessed image will undergo segmentation process. Once the image is segmented using CLAHE method, the labeling of classes takes place. Next, ResNet based classification model will be built by proper training phase. Once the model is created using ResNet, test input images can be provided to attain proper output. At the end of the testing phase, the presented model properly identities the type of DR among the five classes namely normal, mild, severe, proliferate and extremely severe.

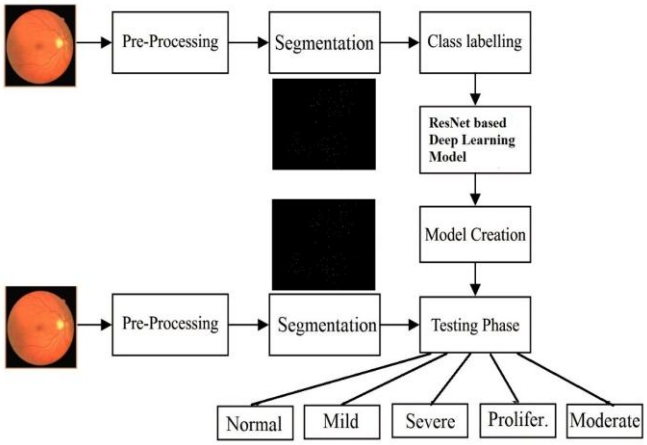


Fig. 1 Overall process of the presented model

A. Contrast Limited Adaptive Histogram Equalization (CLAHE)

In local areas, to prevent additional noise amplification, the improvement is controlled in CLAHE. Numerous valued histograms intensity are estimated through CLAHE, every one of them depends on unique image area, and shares the histogram to prevent additional amplification and intensity values are remapped by employing distributed histograms. For medical imaging, CLAHE was built to improve the low-contrast images. The steps involved in the CLAHE model are given below.

Derive entire inputs: Image, Number of regions in row and column directions, number of bins for the histograms employed in the construction of image transform function (dynamic range), clip limit for contrast limiting (normalized from 0 to 1).

Pre-process the inputs: From the normalized rate, find real clip limit when needed and pad the image prior to partitioning into regions.

Process every contextual area (tile) thus creating mappings by gray level: Derive a single image area, create a region histogram using the certain bin count, clip the histogram by the use of clip limit and generate a mapping (transformation function) for this region.

To collect end CLAHE image, interpolate gray level mappings: Filter out the cluster of four nearby mapping functions, processes image region which partially overlaps with every mapping tiles, extracts individual pixel, applies four mappings to the pixel, and interpolate among the results to attain the output pixel; reiterate over the whole image.

B. ResNet

The highly deep neural networks are difficult to train as they are highly probable to explode or vanish gradients. The activation unit can be fed into a deep network layer to resolve this problem called as skip connection. This forms the foundation for ResNets or residual networks. The image classification result like human level is attained through deep CNN.

In peer-to-peer multi-layer manner, the deep networks derive middle, low and high-level features and the feature "levels" are enriched through the stacked layers counts. The problem of degradation had been uncovered while the deeper network begins to converge. The accuracy is saturated with the increase in network depth and it rapidly reduces. This degradation is not because of embedding numerous layers to deep networks and through overfitting which tends to high training error. The training accuracy deterioration demonstrates that not the entire model is simple to optimize. A framework of deep residual learning is introduced through Microsoft to avoid this issue. It let the layers explicitly suit in residual mapping in spite of hoping each stacked layers to suit an expected underlying mapping. Through feedforward NN, the F(x)+x formulation might be done with shortcut connections. The connections are the one that are skipping numerous layers. Identity mapping is done by shortcut connections and the outputs are embedded towards the stacked layer outputs. There are numerous problems that might be resolved through employing residual network: ResNets are simple to optimize however "plain" networks demonstrate huge training error while depth increases. From highly enhanced depth, ResNets might gain accuracy simply that are superior when compared to prior networks.

Function Classes: Let us assume a function class $\mathcal{F}$ which a certain network framework might reach. There are few parameters set W for entire $f \in \mathcal{F}$ which might be gained by appropriate dataset training. Let us consider the function f^*. If the searching function is in $\mathcal{F}$, there is a good structure. Otherwise, we might try to find few $f_{\mathcal{F}}^*$ which it is superior in $\mathcal{F}$. For example, through resolving the below problem of optimization, we attempt to find it.

$$f_{\mathcal{F}}^* := \underset{f}{argmin} L(X, Y, f) subject to f \in \mathcal{F} \quad (1)$$

It is only reason to consider that when we model varying and highly powerful framework $\mathcal{F}'$. We desire that $f_{\mathcal{F}'}^*$ is 'superior' when compared to $f_{\mathcal{F}}^*$. But, it is not sure to occur when $\mathcal{F} \nsubseteq \mathcal{F}'$, $f_{\mathcal{F}'}^*$ might get worsen. There is a case that we frequently find during implementation that embedding layers only does not create the network highly expressive, it gets modified some times in various ways. The slight abstract terms are demonstrated through the Figure 2.

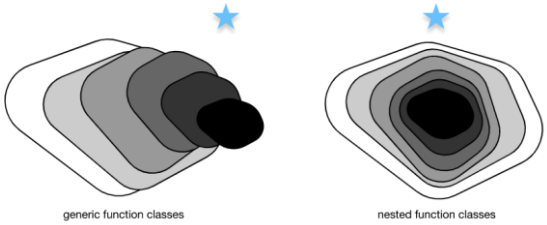


Fig. 2. Left: non-nested function classes; Right: with nested function classes

When a huge function class comprises the small ones, we make sure that enhancing it by network expressive power. Each additional layer must comprise an identity function which acts as a ResNet heart. When we trained the added layer to identity mapping $f(x) = x$, the novel model would be efficient as actual model. To fit the dataset given to training, a novel model might derive superior solution, and it is used to minimize errors during the training process. In a layer, the identity function might be null $f(x) = 0$. They tend to a simple solution which is called as residual block.

Residual Blocks: Let us aim over local neural network as demonstrated in Figure 3. Here, the input is denoted through x. We consider the ideal mapping to derive through f(x) learning to be employed as activation function input. The part that fit in dotted-line box should map f(x) directly. When the certain layer is not needed, this would be difficult rather it retrains input x. The part in dotted-line box in the right image requires to deviation parametrization from identity, as it is returned as x+f(x). Residual mapping is simple to optimize probably. There is a requirement to set f(x) to zero.

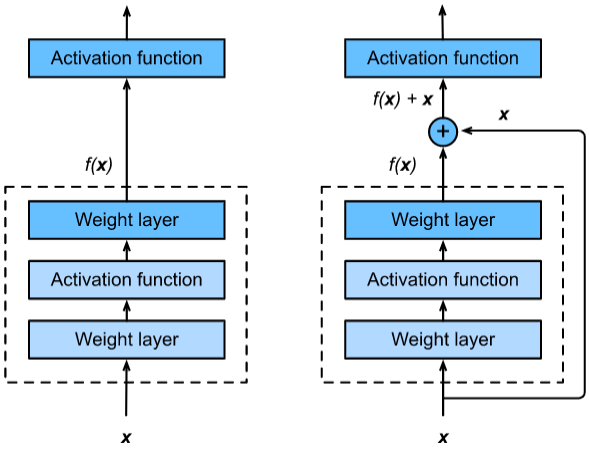


Fig. 3. Difference between regular block (left) and residual block (right)

VGG's are followed by ResNet in the design of 3×3 convolutional layer. With similar output channel counts, the residual block comprises two convolutional layers by the size of 3×3. Through a ReLU activation function and batch normalization layer, every convolutional layer is followed. Prior to the last ReLU activation function, the inputs are added directly and two operations of convolution are skipped. This design type needs two convolutional layers output of similar shape as input as it might be together used. When there is a need to modify the channel count or stride, an extra 1×1 convolutional layer is introduced to change the input into expected shape.

ResNet Model: The primary two ResNet layers are similar as GoogLeNet are defined as: with 64 output channels, 7 × 7convolutional layers and 2 strides subsequent to 3×3 maximum pooling layer with 2 strides are used. The variation is layering of batch normalization subsequently every ResNet convolutional layer. Four blocks are used by GoogLeNet that is made of Inception blocks. But, four modules are used by ResNet, that every one of them employs various residual

blocks with similar output channel counts. The channel count in initial module is similar to the input channelcounts. Therefore, with 2 strides, maximum pooling layer is employed already and it is not needed to minimize the width and height. The channel count is doubled when compared to prior module in the initial residual block every module and width and height are halved.

Towards ResNet, we add entire residual blocks. For every module, two residual blocks are employed. As similar to GoogLeNet, we use the layer of global average pooling subsequent to entirely connected layer output. In every module, there exist four convolutional layers. With the initial and final convolutional and connected layer, there exist 18 total layers. Hence, this is generally called as ResNet-18. In module, through setting various residual blocks and channel counts, various ResNet models are created like deeper 152-layer/ResNet-152. The fundamental ResNet framework is same as of GoogLeNet and ResNet's framework is easy to change. The entire factors result in distributed and rapid ResNet usage. When the channel count enhances, the resolution decreases as like in prior frameworks until wherever entire features are combined by global average pooling layer.

III. PERFORMANCE ANALYSIS

A. Dataset Description

To evaluate the effective segmentation and classification of DR Images, CLAHE and ResNet model is projected and experimentation is done and compared with other existing methods using fundus image dataset. The dataset is derived from Kaggle [21] website which is open source that attempts to build DR detection model.

The eye images of high-resolution are comprised by the dataset and it was graded through trained professionals into five classes as tabulated in table 1. Figure 4 shows the rise of DR from healthy retina towards PDR. In numerous imaging variants, the dataset has RGB images of high resolution upto 35,126 with each having 3500x3000 resolution.

Table I Dataset Description

Class Name	DR Grades	Number of Images
Class 0	Normal	25810
Class 1	Mild	2443
Class 2	Moderate	5291
Class 3	Severe	873
Class 4	Proliferative	708

The labels are given through experts in which it is ranked as DR existence through a scale 0, 1, 2, 3, 4, that starts from no DR, mild, moderate, severe and proliferative DR, correspondingly. The table shows the class names and image counts and grades of DR.

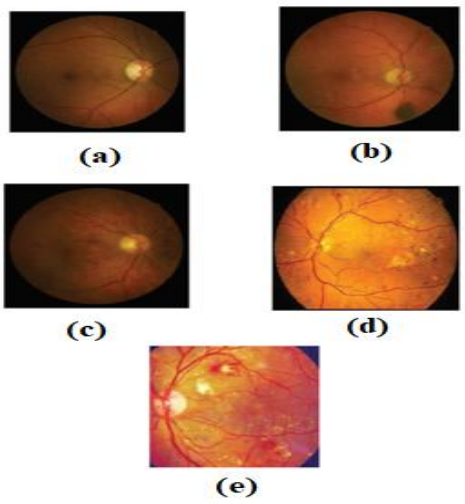


Fig. 4 Stages of DR starting from healthy fundus image(a)Healthy Retina,(b)Mild NPDR,(c)Moderate NPDR,(d)Severe NPDR,(e) PDR.

B. Performance measures

In clinical diagnosis, sensitivity is used to examine the images that precisely recognize the instances with the disease (true positive rate), wherever the specificity is used to examine that it precisely recognizes those who are not acquired of that disease (true negative rate). Accuracy is the percentage of which it classified the instances precisely. Classification accuracy is the ratio of correct predictions to total predictions made. It is often presented as a percentage by multiplying the result by 100. The formulas used to compute sensitivity, specificity and accuracy are given in Eqs. (2)-(4).

$$Sensitivity = \frac{TP}{TP+FP} \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

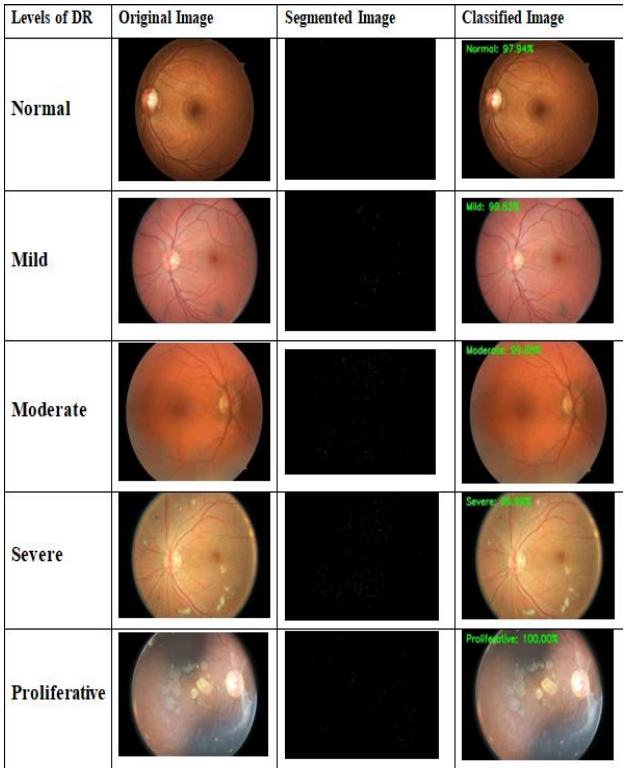
$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

Where TP, TN, FP and FN represent True Positive, True Negative, False Positive and False Negative respectively.

C. Result analysis

Table 2 demonstrates the segmentation and classifier results of the applied DR images. With the various levels of DR such as normal, mild, moderate, severe and proliferative DR, the dataset images are given towards the proposed method of segmentation. The fundus images are segmented using CLAHE method which uses the idea of computing several histograms in distinct image sections by threshold morphology. The segmented images are subjected towards the ResNet classifier to classify the DR level which it is shown in table with classification percentages.

Table II Results of Diabetic Retinopathy



From the experiments, the confusion matrix is derived which is of 5×5 matrix that ranges from zero to four as given in Table 3. With the total number of 500 images with each DR level having 100 images, the confusion matrix of various DR levels is given. The labels of 5×5 matrix are normal, mild, moderate, severe, proliferative.

For determining the classifier efficiency, the derived 5×5 matrix is converted into 2×2 matrix by TP, TN, FP and FN. The derived 2×2 matrix manipulations from confusion matrix are given in table 4. For instance, the 64 instances are detected as normal, whereas 59 out of the total number of instances are detected as proliferative.

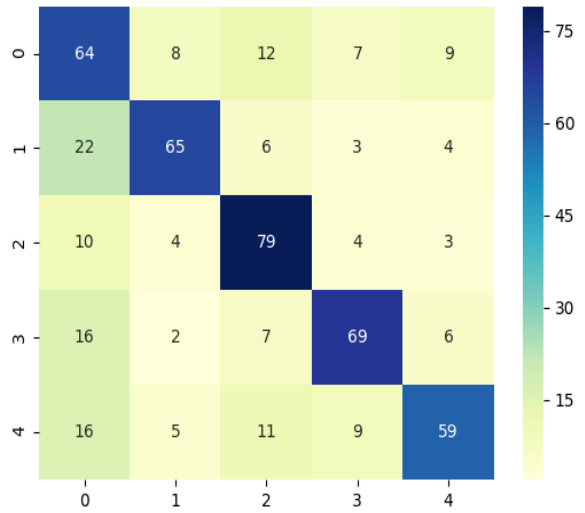


Fig. 5. Confusion Matrix

Table III Confusion Matrix

Input Label	Different Level of Diabetic Retinopathy					Total No. of Images
	Normal	Mild	Moderate	Severe	Proliferative	
Normal	64	8	12	7	9	100
Mild	22	65	6	3	4	100
Moderate	10	4	79	4	3	100
Severe	16	2	7	69	6	100
Proliferative	16	5	11	9	59	100
Total No. of Images	128	84	115	92	81	500

Table IV Manipulations from Confusion Matrix

DR Level	Normal	Mild	Moderate	Severe	Proliferative
TP	64	65	79	69	59
TN	272	271	257	267	277
FP	64	19	36	23	22
FN	36	35	21	31	41

Table V Performance Measures of Test Images with Different DR Levels

Input Grades	Accuracy	Sensitivity	Specificity
Normal	77.06	64.00	80.95
Mild	86.15	65.00	93.45
Moderate	85.36	79.00	87.71
Severe	86.15	69.00	92.06
Proliferative	84.21	59.00	92.64

To evaluate the classifier, it is examined through the performance measures such as sensitivity, specificity and accuracy as shown in Figure 6 and Table 5.

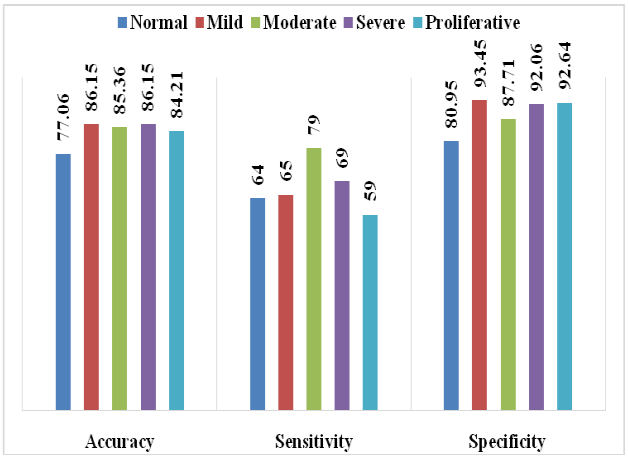


Fig. 6. Comparison of performance over Test Images with Different DR Levels

For normal instances of confusion matrix, the attained accuracy is 77.06% whereas the sensitivity and specificity are 64% and 80.95%. For mild instances, the obtained accurate rate is 86.15% wherever for specificity and sensitivity of the projected classifier attains 65% and 93.45%.

Table VI Performance Measures of Test Images with Various Models

INPUT GRADES	ACCURACY	SENSITIVITY	SPECIFICITY
PROPOSED	83.78	67.20	89.36
VGGNET-S	73.66	33.43	93.98
VGGNET-16	48.13	86.37	29.09
VGGNET-19	82.17	54.51	96.05

For moderate instances of confusion matrix, the attained accuracy is 85.36% whereas the sensitivity and specificity are 79% and 87.71%. The classifier attains accuracy rate of 86.15% for severe instances, and for sensitivity and specificity it offers 69% and 92.06%. For proliferative instances, the accuracy rate of 84.21% is attained through the classifier, whereas the sensitivity and specificity attained is 59% and 92.64%.

The proposed CLAHE-ResNet model is compared with other recently proposed models in order to validate their performance under various metrics like accuracy, sensitivity and specificity as shown in Figure 7 and Table 6. The compared models are VggNet-16, VggNet-19 and VggNet-s. For accuracy, the VggNet-16 is the model that attains the lowest accuracy rate of 48.13%. The models like VggNet-19 and VggNet-s does not show considerable results by obtaining the accuracy rate of 82.17% and 73.66%. The projected model attains the higher accuracy rate of 83.78%. For sensitivity, the lowest accuracy rate of 33.43% is attained by VggNet-s. The projected method attains the sensitivity rate of 67.20%. For specificity, the lowest performance is attained through VggNet-16 whereas the projected method attains 89.36% as specificity rate.

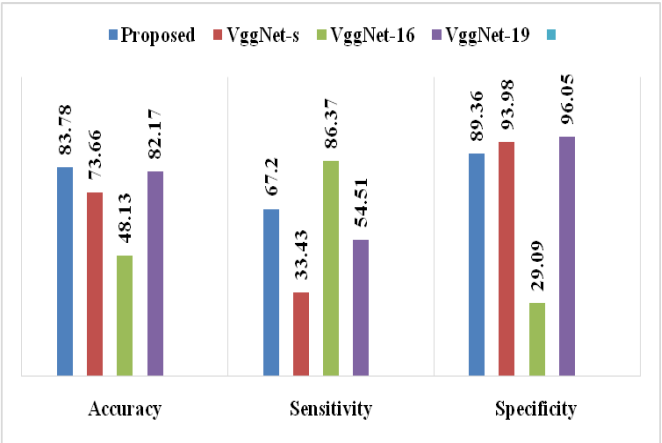


Fig. 7. Comparison of performance over Test Images with various models

IV. CONCLUSION

In this paper, we have presented an efficient image segmentation based classification model to automatically segments and classify the stages of DR. Here, CLAHE is used for segmenting the images place. Next, ResNet based classification model will be built by proper training phase. Once the model is created using ResNet, test input images can be provided to attain proper output. For experimental analysis, the dataset is derived from Kaggle website which is open source platform that attempts to build DR detection model. Finally, a detailed comparative study with the recently presented models also takes place to ensure the betterment of the presented model. The highest classifier performance is attained by the presented model with the maximum accuracy of 83.78, sensitivity of 67.20 and specificity of 89.36 over compared models.

REFERENCES

1. C. P. Wilkinson, F. L. Ferris, R. E. Klein, P. P. Lee, C. D. Agardh, M. Davis, D. Dills, A. Kampik, R. Pararajasegaram, and J. T. Verdaguer, "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," Opthalmology, vol. 110, no. 9, pp. 1677–1682, Sep. 2003.
2. M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "Blood vessel segmentation methodologies in retinal images -A Survey," Comput. Meth. Prog. Bio., vol. 108, pp. 407–433, Mar. 2012.
3. J. Nayak, P. S. Bhat, U. R. Acharya, C. M. Lim, and M. Kagathi, "Automated identification of diabetic retinopathy stages using digital fundus images," J. Med. Syst., vol. 32, pp. 107–115, 2008.
4. Osareh, B. Shadgar, and R. Markham, "A computationalintelligence-based approach for detection of exudates in diabetic retinopathy images," IEEE Trans. Inf. Technol. Biomed., vol. 13, no. 4, pp. 535–545, Jul. 2009.
5. M. U. Akram and S. A. Khan, "Multilayered thresholding-based blood vessel segmentation for screening of diabetic retinopathy," Eng. Comput., vol. 29, pp. 165–173, 2013.
6. P. Bankhead, C. N. Scholfield, J. G. McGeown, and T. M. Curtis, "Fast retinal vessel setection and measurement using wavelets and edge location refinement," PLOS One, vol. 7, no. 3, pp. 1–12, Mar. 2012.
7. J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. v. Ginneken, "Ridge-based vessel segmentation in color images of the retina," IEEE Trans. Med. Imag., vol. 23, no. 4, pp. 501–509, Apr. 2004.

8. M. S. Miri and A. Mahloojifar, "Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction," IEEE Trans. Biomed. Eng., vol. 58, no. 5, pp. 1183–1192, May 2011.
9. U. T. V. Nguyen, A. Bhuiyan, L. A. F. Park, and K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multiscale line detection," Pattern Recogn., vol. 46, pp. 703–715, 2013.
10. S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "DREAM: Diabetic retinopathy analysis using machine learning," IEEE J. Biomed. Health Inform., vol. 18, no. 5, pp. 1717–1728, Sep. 2014.
11. Y. Zhao, L. Rada, K. Chen, S. P. Harding, and Y. Zheng, "Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images," IEEE Trans. Med. Imag., vol. 34, no.9, pp. 1797–1807, Sep. 2015.
12. T. Mapayi, S. Viriri, and J-R. Tapamo, "Comparative study of retinal vessel segmentation based on global thresholding techniques," Comput. Math. Method. M., vol. 2015, pp. 1–15, Nov. 2014.
13. E. Imani, M. Javidi, and H-R. Pourreza, "Improvement of retinal blood vessel detection using morphological component analysis," Comput. Meth. Prog. Bio., vol. 118, pp. 263–279, 2015.
14. P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," IEEE Trans. Med. Imag., vol. 35, no. 11, pp. 2369–2380, Nov. 2016.
15. R. Pires, S. Avila, H. F. Jelinek, J. Wainer, E. Valle, and A. Rocha, "Beyond lesion-based diabetic retinopathy: a direct approach for retinal," IEEE J. Biomed, Health Inform., vol. 21, no. 1, pp. 193–200, Jan. 2017.
16. Wang Z , Yang J . Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation. arXiv preprint arXiv:170310757 2017
17. Wang S , Yin Y , Cao G , Wei B , Zheng Y , Yang G . Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. Neurocomputing 2015; 149:708–17
18. KrizhevskyA ,Sutskever I , Hinton GE . Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems; 2012. p. 1097–105.
19. Szegedy C , Liu W , Jia Y , Sermanet P , Reed S , Anguelov D , Erhan D , Vanhoucke V , Rabinovich A , et al. Going deeper with convolutions. Cvpr; 2015.
20. Parkhi OM, Vedaldi A, Zisserman A, et al. Deep face recognition. In: BMVC, vol. 1; 2015. p. 6.
21. https://www.kaggle.com/c/diabetic-retinopathy-detection/data