Advances in Parallel Computing Technologies and Applications
D.J. Hemanth et al. (Eds.)

© 2021 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/APC210120

A Contemporary Method on Feature Selection and Classification Using Multi-Model Deep Learning Technique for Identifying Diabetic Retinopathy

Meenakshi G [a,1] and Thailambal G [b]
[a] *Department of Computer Science, JBAS college for women (Autonomous), Chennai, Tamil Nadu*
[b] *Department of Computer Science, School of Computing Sciences, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu*

Abstract. Diabetic retinopathy is one of the leading reason for preventable blindness in the world. 10 -18 % of diabetic people having diabetic retinopathy. The feature selection and classification is a vital task to find the seriousness of the diabetic retinopathy. The different researchers introduced different techniques to extract the features and classification of diabetic retinopathy images. The deep learning is one of the essential methods to extract the features. Most of the previous techniques are extracted information's with the help of texture and extracted the whole image feature data. Some feature missed and thereby the accuracy is significantly less. Hence a proposed new technique called FRCNN (Fast Region-based Convolution Network) and Nearest Neighbour (NN) algorithm used to extract the features and classifications. The proposed method yields better accuracy (96%), sensitivity (98%) and specificity (97%) compared to the previous methods. The implementations Messidor Dataset is used for training and testing

Keywords. Diabetic Retinopathy, Features, Classification, Nearest Neighbour, Deep Learning, FRCNN (Fast Region-based Convolutional Network).

1. Introduction

Diabetic retinopathy (DR) is one of the leading reason for preventable blindness in the world. As per the survey [1–3] 210 million people are having diabetics and around 10-18 % of people having diabetic retinopathy. The DR patients have affected the blood vessels of light sensitivity parts or tissue at the retina. Initially, the patents has no symptoms, but in the cases of type 1 and type 2 diabetics automatically affect the eye and eye diseases getting started simultaneously. The main symptoms of DR are blurred vision, impaired colour, fluctuating vision, dark vision, vision loss and empty areas in the vision etc. The DR occurred in the age of 25 to 74 years of ages. As per the clinical features,

[1] Meenakshi G, Department of Computer Science, JBAS college for women (Autonomous), Chennai.
E-mail: meenakshi2020research@gmail.com.

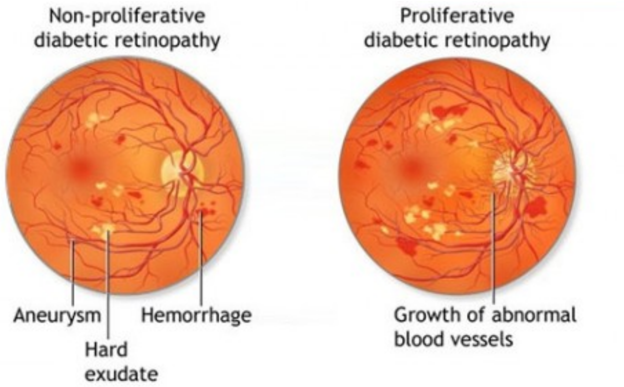

Figure 1. Representation of NPDR and PDR.

the DR is classified into three types such as [4] Non-proliferative diabetic retinopathy (NPDR), Proliferative diabetic retinopathy (PDR) and macular edema. The NPDR and PDR representation is shown in Figure 1, and sub-classification types shown in table

Table 1. Types, meaning and sub-classifications of DR

S.No	Types	Meaning	Sub- Classifications
1.	NPDR	The initial stage of diabetic retinopathy. it damages the blood vessels in the retina and begin to start fluid into the eye.	1. Mild NPDR. 2. Moderate NPDR 3. Severe NPDR 4. Very Severe NPDR.
2.	PDR	Close to blood vessels in the retina and stop the blood flow to retina. The blood vessels leak the blood and loss the eye site and time to time check needed.	1. PDR with neovascularisation 2. PDR with neovascularisation of the disc
3.	Macular Edema	Accumulation and abnormal leakage of fluid from damaged vessels in the retina.	—

The Table 1 shown the various types, meaning and sub- classifications of DR. This classifications and sub-classifications are identified based on the symptoms. The author of [5] presented the various classification based on the affecting status. Based on the affecting status, graded as R0, R1, R2 and R3. The R0 means not affected, R1 means mild affected (NPDR), R2 means severe affected (NPDR), R3 means most severe affected (PDR). So classification features of image is very important in DR. The different techniques, algorithms and methodologies are introduced to finding the features. The machine learning techniques are used to find the features of the images, one should note that Machine learning is applied in variety of application such as [6] cost control, soil environment preservation, Medical etc,. but, all these benefits do come with certain limitations. But the comparison parameters such as sensitivity (Sen), specificity (Spe) , accuracy (Acc) of the feature's predictions are not yet to the mark. So, features selection and classification methods are needed for further improvement in terms of all parameters. The previous works such as SVM (Sen-80%, Spe- 86% and Acc- 83%), SVM+BPSO (Sen-94%, Spe-98% and Acc-96%) [7] and another recent work [8] received very less accuracy compare to the SVM + BPSO model.

New techniques and methodologies are required to improve the accuracy, sensitivity and specificity. In this work, planning new deep learning-based techniques to improve the prediction of features and increase the accuracy of prediction features. The deep learning is one of the main emerging techniques. The main advantages of deep learning over machine learning and artificial intelligence is, searching the features on its own and take multiple features that combine, correlate with other relevant features to fast learning. So, it, produce better results compared to the other methodologies. The main contribution of this work is

1. Introduced multi-model techniques to predict features with the help of deep learning
2. Used relevant and morphological are used to better predictions of features.
3. The accuracy, sensitivity and specificity are increased compared to the other recent methods such as SVM, SVM+BPSO [7] and colour histogram filter method [8].

2. Related Work

The different researchers introduced, different methods to predict the features of DR. In this section some of the works related to support for improving the features prediction. The authors of [9] presented automated detection of new vessels in the retina. This work is supported to dual classification techniques. The Local morphology features are used to measure the DR features. The authors of [10] presented a method for finding DR macular edema features with the help of deep learning technique. The FRCNN method with fuzzy k-means clustering used for features finding. The authors of [11] present particle swarm optimization method for finding the features of DR. In this work the selected features are again classified with the help of neural network. The accuracy of features prediction is 76.11%.

The authors of [12] introduced genetic algorithm with dual classification method for automated detection of PDR. The SVM classification is used with genetic algorithm for finding features. The authors of [13] presented conventional neural network method for screening various stages of DR. The different stages such as NPR and PDR in different stages are measured. The authors of [14] introduced deep learning techniques for features selection. This method used deep belief network for classification and MGS-ROA method is used for features selection. The authors of [7, 8] introduced multi-model or hybrid method for features in DR. These two methods given the better results compared to the other methods. The authors of [15–17] introduced various deep learning methods to select the features of DR. These deep learning methods are mostly supported to take the better decision. The authors of [18] present texture-based features extraction with the help of nearest pixel in the large areas. The authors of [19] present the morphological based features in the DR. With the help of morphological features easily find the relevant features also. The most of the previous work is the texture and morphological features are extracted separately and deep learning-based work given better results compared to the other artificial intelligence and machine learning.

3. Materials and Metho

The MESSIDOR Dataset [20] used for features classification, training and testing. The MESSIDOR Dataset consists of total 1200 fundus colour images. The 300 images in the dataset are used for training and 100 images in the dataset are used for testing. The training data is used for validation of results and testing data is used for verifications of the images.

3.1. FRCNN + NN Method

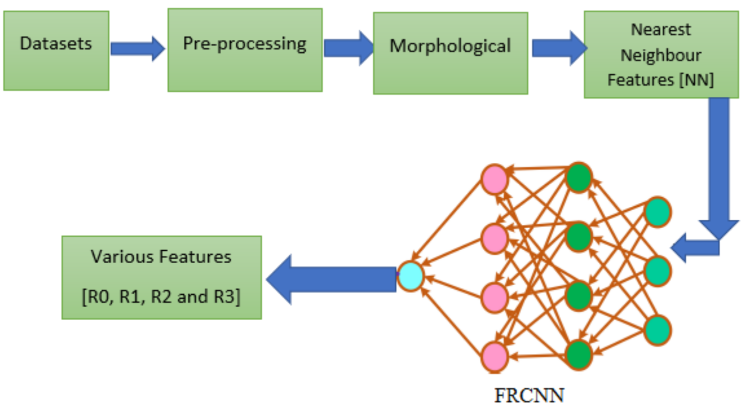

Figure 2. Proposed FRCNN + NN method.

The DR image finding FRCNN + NN method proposed for feature extraction and classification of the images. The proposed consists of four parts such as i. pre- processing ii. Morphological iii. Nearest Neighbour (NN)features extraction and iv. FRCNN learning. The pre-processing is used to check the quality of the images and extraction of useful images. The morphological features used to remove the imperfections and accounting of structure of features. The nearest neighbour features used to find the nearest features of images and classification. The FRCNN (Fast Region-based Convolutional Network) learning is used to find the extract features in the fastest way. The FRCNN method is used to give the output in fastest way.

3.1.1. Preprocessing of DR images

The pre-processing of DR images [21,22] having different steps such as color space, Spatial normalization, Region of Interest (ROI) extraction, Illumination correction, contrast enhancement and Vessels extraction etc. The entire process of preprocessing of images shown in the Figure 3. Using the pre-processing all the unwanted issues such as color, location and ROI extraction and vessels issues all are extracted.

3.1.2. Morphological Extraction

Morphological extraction is used to denoise the original images, identify the improper choice, length of the features, size and shape of the images to predict the features. The

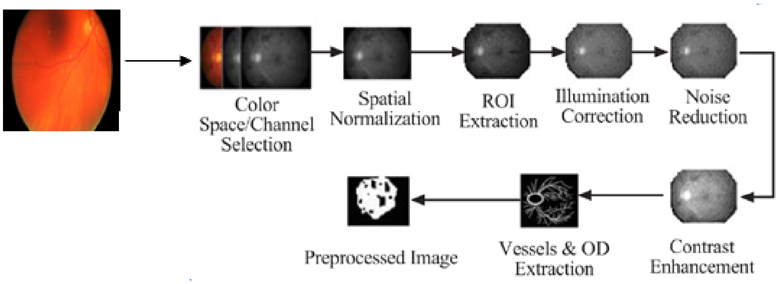

Figure 3. Pre-processing of DR images.

following function such as erosion Eq. (1), dilation Eq. (2), opening Eq. (3), closing Eq. (4) and differences Eq. (5) are used to extract the better resolutions of the images.

$$(f \ominus g)(n) = min[f(n+m) - g(m)) \tag{1}$$

Where m = 0,1,2,3,. . . ,M-1, n = 0,1,2,,..,N-1

$$(f \oplus g)(n) = min[f(n+m) + g(m)) \tag{2}$$

Where m = 0,1,2,3,. . . ,M-1, n = 0,1,2,,..,N-1

$$(f \circ g)(n) = (f \ominus g \oplus g)(n), n = 0,1,2,3\ldots N-1 \tag{3}$$

$$(f \cdot g)(n) = (f \oplus g \ominus g)(n), n = 0,1,2,3\ldots N-1 \tag{4}$$

$$((f \cdot g) - (f \circ g)(n)) = (f \oplus g \ominus g) - (f \ominus g \oplus g)(n) \tag{5}$$

f(n)- original function, f(m)- collected information's, $\ominus$- operation of erosion, $\oplus$- operation of dilation.

3.1.3. Nearest Neighbour (NN)

The NN features selection or classification method is a supervised algorithm. Using this NN, the nearest features of pixel, shape, structure, PDR blood vessels, NPDR blood vessels easily can classify and find the features. The Euclidian distance is used to find the distance between one pixel to another pixel. The distance metrics and equation shown in the Eq. (6).

$$D(L1, L2) = \sum_{n=1}^{\infty} (P \vee LP1 - LP2) \tag{6}$$

The R-CNN is the base for faster R-CNN. It is used to find the selective features in the conventional network. This method combines all the similar pixels and morphological features. In this FRCNN the different rectangular areas feature also combined with multiple regions. In the proposed work FRCNN and NN features also used to getting the final output. At the time of computation of FRCNN the previous regions or previous steps also used to extract the features simultaneously. The pooled conventional neural networks filter and extract the selective features form the various features. So, the accuracy of the features also increased automatically. The overall process of features extraction combined with NN and morphological features shown in the Figure 2. The overall step by step process shown below:

1. Initialize the dataset
2. Perform the pre-processing steps
3. Find the morphological features with the help of equation 1-6
4. Nearest features are extracted using NN (Equation -7)
5. Training and testing performed using FRCNN.
6. Various features of grading performed using accuracy, sensitivity and specificity.

4. Results and Discussion

The experimental purpose MESSIDOR dataset is used [20]. The MESSIDOR datasets used for training and testing using FRCNN deep learning. The MESSIDOR dataset having 1200 images and the images are grouped into two categories such as DME and DR. In this work the image is used for only DR classification and features selection. The images were captured using high resolution images such as 1440*960, 2240*1488 and 2404 *1536.

Table 2. Comparison of proposed work with various other methods

S. No	Methods	Accuracy	Sensitivity	Specificity
1	SVM [9]	83%	80%	86%
2	SVM+BPSO [9], 2019	94%	98%	96%
3	MGS-ROA -DBN [16], 2020	93%	86%	95%
4	Proposed Work & FRCNN + NN	96%	98%	97%

The proposed work implementation, 300 images are used for training and 100 images are used for testing. The proposed work FRCNN+NN method used for training and testing. The proposed work implemented and performed in the different iterations. Each iterations the number of testing and training images are same. The proposed work evaluated with the help of three parameters such as accuracy, sensitivity and specificity. The previous methods such as SVM [7], SVM+BPSO [7], DT (J8) + K -NN [8] and MGS-ROA -DBN [14] also used same parameters for evaluations. These parameters are performed in various iterations, and results are consolidated. The proposed work consolidated data and comparison with various methods shown in the Table 2.

The Figure 4 shown the accuracy of proposed work and comparison with various other method. The proposed work produced better result such as 96% and compared with other methods it produces better results because two attributes are used to increase

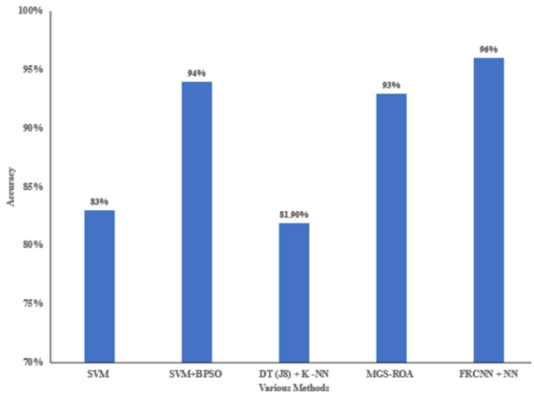

Figure 4. Accuracy of features selection.

the accuracy. The nearest features are selected using NN and FRCNN is increase the accuracy. The validation of accuracy is defined as shown in the Eq. (7).

$$Accuracy = \frac{TP+TN}{TN+TP+FN+FP} \tag{7}$$

Where TP -true positive, TN - true negative, TP – Ture positive and FP- False Positive. Similarly, the validation of sensitivity and specificity shown in the Eq. (8) and Eq. (9).

$$Sensitivity = \frac{TP}{FN+TP} \tag{8}$$

$$Specificity = \frac{TN}{FP+TN} \tag{9}$$

The sensitivity is used to find the correctly matched features from the datasets. The specificity is used to find the wrongly matched images from the datasets. The prediction of sensitivity shown in the Figure 5 and the experiment is performed in the different iterations and corresponding average values are plotted in the Figure 3 and Table 2.

Similarly, the specificity prediction in the Figure 6 and the average data mentioned in the Table 3. With the help of accuracy, sensitivity and specificity the four grading data is predicted such as R0, R1, R2, and R3. The predicted grading data shown in the Table 2. The grading data predicted from the 100 testing images.

Table 3. Comparison of proposed work with various other methods

Iterations/ Grading	R0 Not affected	R1 Mild affected	R2 Severe affected	R3 Most severe affected
Iteration -1	58	22	14	6
Iteration -2	55	24	16	5
Iteration -3	58	22	14	6
Iteration -4	58	22	14	6
Iteration -5	57	23	14	6

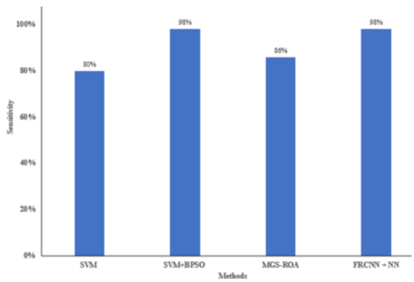

Figure 5. Sensitivity of features selection

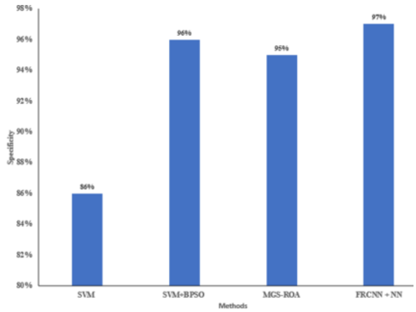

Figure 6. Specificity of features selection

5. Conclusion

Diabetic retinopathy is an important source of vision damage in the world. The different indications are used to identify the vision loss in the retina. In the DR computational features selection and classification is important to identify the symptoms. Based on the features and classifications the input images are categories into no DR, moderate DR, mild DR, and severe condition DR. In this work proposed multi-model technique FCNN + NN used for classifications and features selection. The FCNN+ NN method consists of pre-processing, morphological, NN and FCNN. The proposed work produces better results in the I age with the help of morphological features. The proposed work additionally used NN classification to find the nearest features in the retina. The FCNN is used to find the better accuracy compared to other methods. The proposed method yields better accuracy (96%), sensitivity (98%) and specificity (97%) compared to the previous methods. The future work, the performance can be improved in terms of features selection and prediction of the DR.

References

[1] Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes care. 2004 May 1;27(5):1047-53.

[2] Day C. The rising tide of type 2 diabetes. The British Journal of Diabetes & Vascular Disease. 2001 Aug;1(1):37-43.

[3] Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes research and clinical practice. 2010 Jan 1;87(1):4-14.

[4] Fraser CE, D'Amico DJ, Trobe J. Diabetic retinopathy: Classification and clinical features. UpToDate. Waltham, MA: UpToDate Inc. 2014.

[5] Antal B, Hajdu A. An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. IEEE Transactions on biomedical engineering. 2012 Apr 3;59(6):1720-6.

[6] Luo Y. Environmental cost control of coal industry based on cloud computing and machine learning. Arabian Journal of Geosciences. 2021 Jun;14(12):1-6.

[7] Chaurasiya RK, Khan MI, Karanjgaokar D, Prasanna BK. BPSO-based feature selection for precise class labeling of Diabetic Retinopathy images. In Advanced Engineering Optimization Through Intelligent Techniques 2020 (pp. 253-264). Springer, Singapore.

[8] Vijayan T, Sangeetha M, Kumaravel A, Karthik B. Feature Selection for Simple Color Histogram Filter based on Retinal Fundus Images for Diabetic Retinopathy Recognition. IETE Journal of Research. 2020 Nov 18:1-8.

[9] Welikala RA, Fraz MM, Dehmeshki J, Hoppe A, Tah V, Mann S, Williamson TH, Barman SA. Genetic algorithm based feature selection combined with dual classification for the automated detection of proliferative diabetic retinopathy. Computerized Medical Imaging and Graphics. 2015 Jul 1;43:64-77.

[10] Nazir T, Irtaza A, Javed A, Malik H, Hussain D, Naqvi RA. Retinal image analysis for diabetes-based eye disease detection using deep learning. Applied Sciences. 2020 Jan;10(6185):1-21.

[11] Asti H, Toni A, Sari S, Hikmah AB. Feature Selection of Diabetic Retinopathy Disease Using Particle Swarm Optimization and Neural Network.In 2018 6th International Conference on Cyber and IT Service Management (CITSM) 2018 Aug 7 (pp. 1-4). IEEE.

[12] Welikala RA, Fraz MM, Dehmeshki J, Hoppe A, Tah V, Mann S, Williamson TH, Barman SA. Genetic algorithm based feature selection combined with dual classification for the automated detection of proliferative diabetic retinopathy. Computerized Medical Imaging and Graphics. 2015 Jul 1;43:64-77.

[13] Shaban M, Ogur Z, Mahmoud A, Switala A, Shalaby A, Abu Khalifeh H, Ghazal M, Fraiwan L, Giridharan G, Sandhu H, El-Baz AS. A convolutional neural network for the screening and staging of diabetic retinopathy. Plos one. 2020 Jun 22;15(6):1-13.

[14] Jadhav AS, Patil PB, Biradar S. Optimal feature selection-based diabetic retinopathy detection using improved rider optimization algorithm enabled with deep learning. Evolutionary Intelligence. 2020 Apr 9:1-8.

[15] Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama. 2016 Dec 13;316(22):2402-10.

[16] Lu D, Heisler M, Lee S, Ding GW, Navajas E, Sarunic MV, Beg MF. Deep-learning based multiclass retinal fluid segmentation and detection in optical coherence tomography images using a fully convolutional neural network. Medical image analysis. 2019 May 1;54:100-10.

[17] Ting DS, Cheung CY, Lim G, Tan GS, Quang ND, Gan A, Hamzah H, Garcia-Franco R, San Yeo IY, Lee SY, Wong EY. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. Jama. 2017 Dec 12;318(22):2211-23.

[18] Du N, Li Y. Automated identification of diabetic retinopathy stages using support vector machine. In-Proceedings of the 32nd Chinese Control Conference 2013 Jul 26 (pp. 3882-3886). IEEE.

[19] Adhi M, Brewer E, Waheed NK, Duker JS. Analysis of morphological features and vascular layers of choroid in diabetic retinopathy using spectral-domain optical coherence tomography. JAMA ophthalmology. 2013 Oct 1;131(10):1267-74.

[20] Decencière E, Zhang X, Cazuguel G, Lay B, Cochener B, Trone C, Gain P, Ordonez R, Massin P, Erginay A, Charton B. Feedback on a publicly distributed image database: the Messidor database. Image Analysis & Stereology. 2014 Aug 26;33(3):231-4.

[21] Ramasubramanian B, Selvaperumal S. A comprehensive review on various preprocessing methods in detecting diabetic retinopathy. In2016 international conference on communication and signal processing (ICCSP) 2016 Apr 6 (pp. 0642-0646). IEEE.

[22] Ashraf MN, Hussain M, Habib Z. Review of Various Tasks Performed in the Preprocessing Phase of a Diabetic Retinopathy Diagnosis System. Current medical imaging. 2020 May 1;16(4):397-426.