Indonesian Journal of Electrical Engineering and Computer Science
Vol. 30, No. 1, April 2023, pp. 81~88
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v30.i1.pp81-88

An efficient machine learning approach for classification of diabetic retinopathy stages

Srilaxmi Dasari[1], Boo Poonguzhali[1], Manjula Sri Rayudu[2]
[1]Department of Electronics and Instrumentation Engineering, Faculty of Engineering and Technology, Annamalai University, Annamalai Nagar, India
[2]Department of Electronics and Instrumentation Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India

Article Info

Article history:

Received Jul 11, 2022
Revised Dec 9, 2022
Accepted Dec 16, 2022

Keywords:

Diabetic retinopathy
EML-DRGC
Machine learning
Multi support vector machines classification
Retinal fundus images

ABSTRACT

Diabetic retinopathy (DR) the prime cause of blindness, develops when glucose levels rise, causing retinal damage. DR can be prevented if the illness is detected early. As a result, early grading, categorization, and diagnosis of DR can help diabetic patients avoid visual loss. Several system methods assist in the classification of DR using high-performance criteria. This work proposes an efficient system-based DR classification. The purpose of efficient machine learning dabetic retinopatyy grading classification (EML-DRGC) design is to recognize DR impulsively with highest accuracy. The proposed technique employs preprocessing methods such as employing the Gaussian filtering approach for removing noise present in retinal fundus images. The segmentation process is followed using K-means segmentation algorithm which is used for segmenting the region of interest (ROI) from background. Moreover, Feature extraction process is done by using gray level co-occurrence matrix (GLCM) in which features are extracted bycapturing the image's visual content and features from acceerated segment test (FAST) design is used as extractor of features. Finally, multi support vector machine is utilized as classifier for detecting severity levels of DR. Performance metrics such as accuracy of 98.38% and specificity of 98.34% are obtained which are superior to existing designs.

Corresponding Author:

Srilaxmi Dasari
Department of Electronics and Instrumentation Engineering, Faculty of Engineering and Technology
Annamalai University
Annamalai Nagar, Chidambaram, India
Email: srilaxmi_d@vnrvjiet.in

1. INTRODUCTION

Diabetic retinopathy (DR) affects both diabetic patients with type 1 and type 2 diabetes. The retinopathy is a pathology of retina. DR caused by persistent diabetes is a group of abnormalities in retina. In many countries, DR is principal source of sightlessness. 80% of patients are unaware of the eye disease when they are first visiting the doctor due to which condition becomes dangerous. So, most cases are due to delays in treatment. On a severity basis, it is divided into 5 levels such as mild, moderate, severe, proliferative DR, and non-proliferative diabetic retinopathy (NPDR). The early stage of DR, NPDR in which retinal vessels weaken producing micro aneurysms which are a kind of microscopic lump to bulge out of the vessels side. Symptoms may be very mild in this stage. Blood arteries get clogged and can only transfer a little amount of blood. Because some areas of the body become deprived of oxygen and transmit data to the retina to form new blood vessels, it is critical for diabetes patients to have frequent checks in order to receive early diagnosis and

treatment. As DR progresses, new blood vessels begin to develop through into retina and into the vitreous, a gel-like fluid that fills the cavity of the eye. proliferative diabetic retinopathy (PDR) is a much more advanced form of DR. These blood vessels are thin, and they may leak blood into the macula, causing vision issues. As the retina is damaged, scar forms, and pressure builds up, inflicting damage to the optic nerve. Early treatment and diagnosis are very much useful for preventing DR in patients [1]-[4]. The physical approach of DR screening used by eye doctors is insufficient to screen the rapidly growing population of diabetic individuals at risk of vision loss. This manual approach takes a longer time for screening diabetic patients. As a result, automated process of classification of diabetic retinopathy severity prediction enhances the diagnosis of DR effectively [5]-[10]. The retinal Fundus images are the most important images for studying DR, which is now under examination. These are mostly used for the analysis of DR disease [11], [12]. Various datasets are Messidor, Drive, Diaretdb1, Kaggle have been used for these observations. Moreover, when compared to alternate methods the proposed technique based on maximum accuracy, specificity obtains improved classification compared to existing methods [13]-[15]. A comparison study of texture features based on the distribution of features were discussed in [16], [17]. Investigations of texture features associated with gray-level co-occurrence matrix (GLCM) statistical parameters are dealt detailly in [18]. Some other methods of classification using different classifiers are dealt in [19]-[22] classification process by Inception V3 is done for identifying DR in [23]. Anomaly detection by 2 variants of hybrid ML Techniques is clearly explained in [24]. Alex support vector machine (SVM) fused HC design built for classifying type of disease was explained in detail in [25]. Classification algorithm based on spatial U-Net is proposed for quantification of AVR in retinal images in [26].

2. RESEARCH METHOD

The complete flow of efficient machine learning dabetic retinopatyy grading classification **(**EML-DRGC) design is shown in Figure 1. According to diagram, EML-DRC design includes a Gaussian filter based preprocessing approach, K-means clustering-form of segmentation of image, and features from acceerated segment test (FAST) design-based image GLCM Multi SVM-based classification and feature extraction. The following sections details the operation of each module subsections.

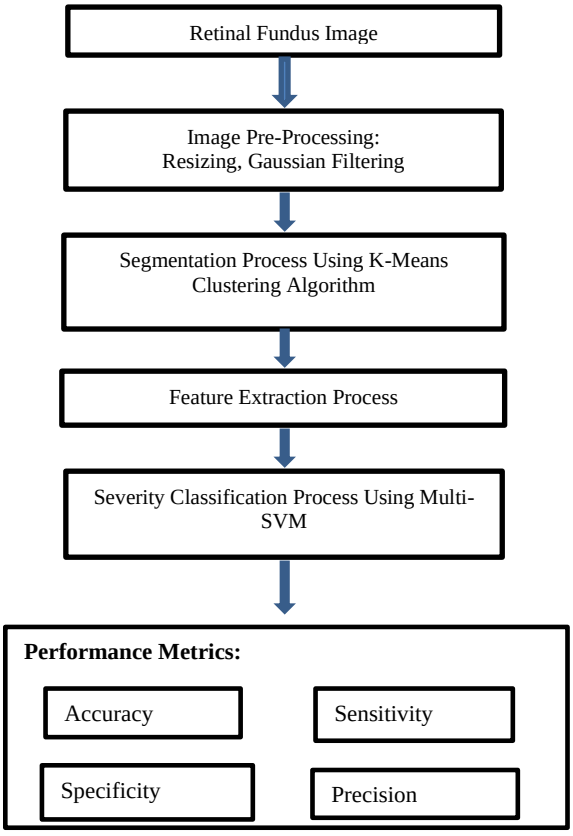


Figure 1. Block diagram of proposed methodology

2.1. Dataset used

The data for the proposed model is a dataset of images taken from Messidor funded by French Ministry of Research and defense. Messidor database consists of 1200 images which are classified into 4 categories. The proposed model uses 20% for testing and 80% for training Figure 2 shows the sample of retinal fundus images used.

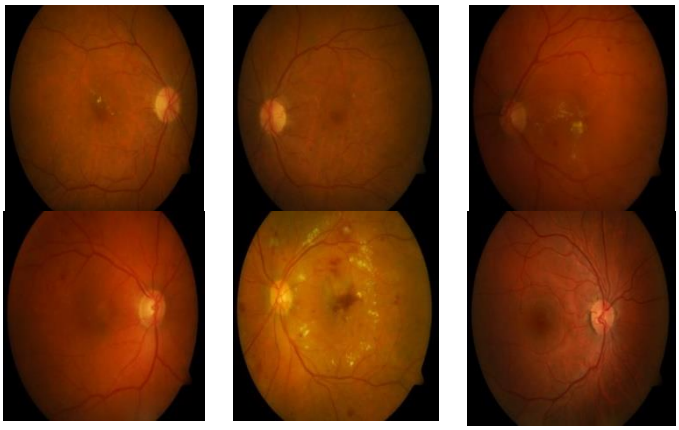


Figure 2. Retinal fundus images

2.2. Gaussian filter preprocessing approach

Images are blurred and noise is removed using Gaussian filtering. In Gaussian filter preprocess approach every pixel is processed, processing distinct colour channels by averaging the pixel value of each pixel according to its convolutional location kernel. The original image will be blurred. 2D Gaussian filter is used for smoothing which is approached by functions of convoltuion.

The Gaussian function in one dimension is:

$$Gx) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-}\left(\frac{x^2}{2\sigma^2}\right) \quad (1)$$

the Gaussian function in two dimensions is:

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

where in $sigma$ represents the 2-D distribution's standard deviation as a point spread function which is achieved by convolution. (x, y) demonstrates the dimensions of the window. With increasing distance from the kernel's centre, the kernel coefficients decrease. The weighting of central pixels is higher than that of peripheral pixels. A broader peak is produced by larger values of greater blurring.

2.3. K-means clustering based image segmentation

Classification of image into different groups is image segmentation the digital image is partitioned into different segments containing super pixels with same attributes. K-means clustering, an unsupervised architecture is the popular method in segmentation of image. Clustering is a process of identifying clusters in the dataset. The goal is to find particular groups based on data similarity, with K being the total number of groups. K-Means clustering aims to lower the sum of squared distances between all points in the cluster and the cluster's centre.

$$J = \sum_{j=1}^{K}.1 \sum_{i=1}^{n} 1\,(x - c)$$

where c is no. of cluster centres.

K-means clustering algorithm: i) initialize K; ii) K points has to be selected randomly; iii) each data point should be assigned to the centroid that is closest to it, resulting in K clusters; iv) place the new cluster centroid after calculation; v) rename the nearest centroid for each data point; and vi) if any reassignment, proceed to step 4, otherwise the design is complete.

2.4. GLCM, FAST design based feature extraction

The gray-level co-occurrence matrix is a statistical texture assessment method that considers the spatial interaction of pixels (GLCM). The output of GLCM element is the sum of the number of times the pixel with value i appeared in the specified spatial relationship to a pixel with value j in the input retinal fundus picture (i, j). The number of grey levels in the image determines the size of the GLCM. The GLCM routines calculate how often pairs of pixels with specified values appear in a picture, generate a GLCM, and then extract statistical measures from it. Statistics about texture features of an image calculated are:

- Contrast: the GLCM local fluctuations.
- Correlation: the possibility of the given pixel combinations matching together.
- Energy: provides information of the summation of squared parts in GLCM.
- Homogenity: the distance between the GLCM diagonal and the distribution of characteristics in the GLCM.
- FAST design-based feature extraction: FAST design finds the corners and returns the object of corner points. The feature points recognized in a 2-Dimensional input picture are stored in this object. The feature points are found using the features from the FAST algorithm using the detect FAST features function.

2.5. Multi-SVM based classification

Multi-label SVM assigns labels to objects using support vector machines, with labels generated out of a finite collection of tuples. Breaking down every single multi-label issue into several 2-class tasks seems to be the most prevalent way. The following are some examples of common ways for reaching such a reduction: In the one-versus-all example, new examples are categorized using a winner-takes-all strategy, wherein the classification with the maximum final output assigns the class.

3. RESULTS AND DISCUSSION

The DIARETDB1 dataset [27] is also used for the suggested design's experimental validation. The collection contains pictures with variying degrees of severity of DR, such as different stages. Stage 1 represents mild DR, stage 2 represents moderate DR, and stage 3 represents severe DR. Table 1 and Figures 3-7 depicts computation of EML-DRC design on sample images by observations of outcomes after experimentation on dataset. According to the EML-DRC design, DR is divided into five phases. With a specificity of 99.90% and 99.93% accuracy, EML-DRC design identified pictures as 'normal.' Similarly, the samples were classed as 'Stage 1' by the EML-DRC design, which had a specificity of 99.3% and 99.35% of accuracy. The samples were also classed as 'Stage 2' by the EML-DRC design, which had a specificity of 98.21% and 98.22% of accuracy. The EML-DRC approach, on the other hand, categorized the samples as 'Stage 3' with a specificity of 95.98% and 96.03% of accuracy. Finally, the suggested EML-DRC approach has average of specificity of 98.34% and 98.38% of accuracy.

Table 1. EML-DRC design performance results

Severity level	Specificity	Accuracy
Normal	99.90	99.93
Stage1	99.30	99.35
Stage2	98.21	98.22
Stage3	95.98	96.03
Average	98.34	98.38

The complete study of the EML-DRC design takes place in different measures with existing methods in Table 2 [28], [29]. From results the proposed EML-DRC design has secured highest accuracy and specificity than existing methods. Retinal fundus images are processed to safeguard the retina blood vessels. Early diagnose using systematic approach will add support to diabetic patients. Detecting the DR at acute stages is a serious concern. This approach provides an automated superior screening system to diagnose the DR in advance. The experimental results of percentage of severity level are authenticated with ophthalmologist. The experimental results for image ‘1’ are corelated with ophthalmologist validation. The experimental results for Image ‘2’ and image ‘3’ are deviated 4% and images ‘5’ and ‘6’ are proximity with ophthalmologist findings. All the results show meticulously estimated with efficient system approach-based classification. Table 3 depicts comparison of percentage of severity level using image processing and experts for sample images.

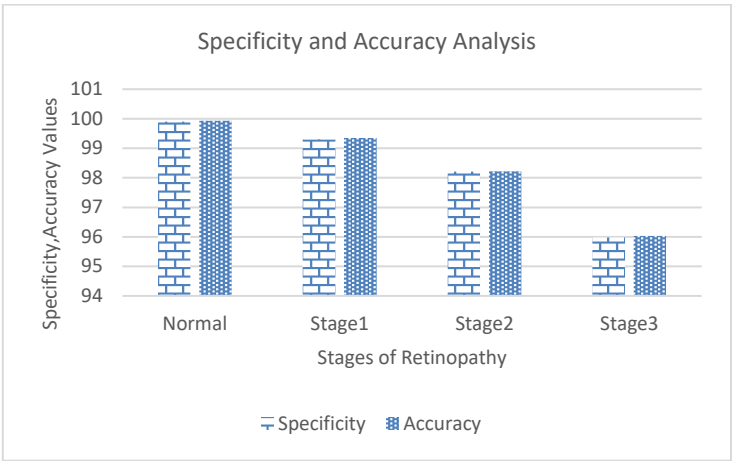


Figure 3. Specificity, accuracy study of EML-DRGC design

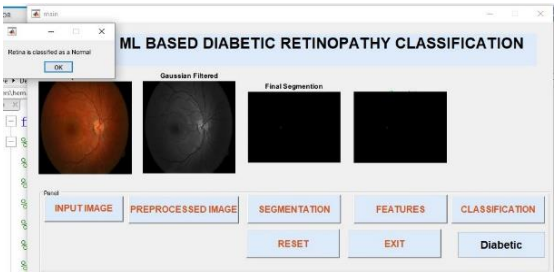


Figure 4. Normal stage output of EML-DRC design

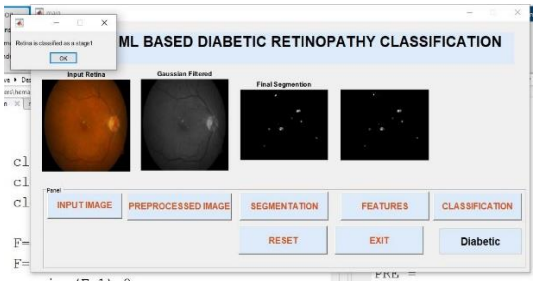


Figure 5. Stage 1 output of EML-DRC design

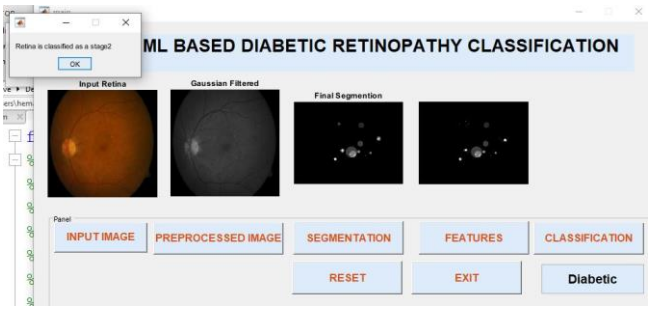


Figure 6. Stage 2 output of EML-DRC design

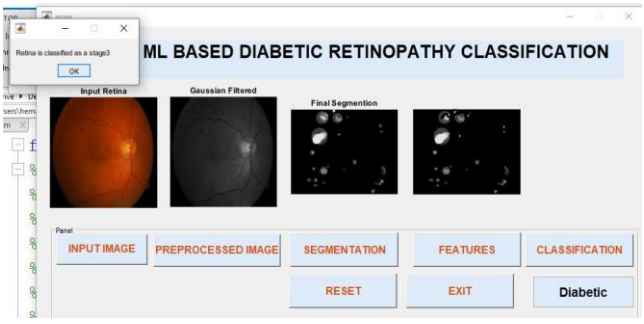
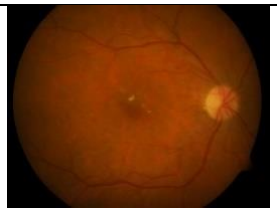
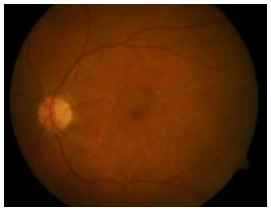
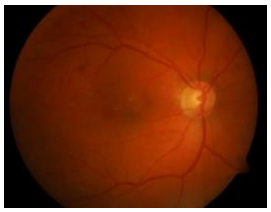
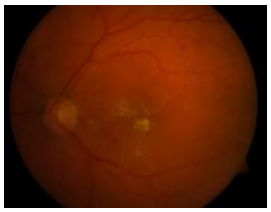
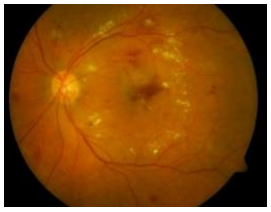
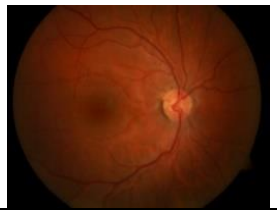


Figure 7. Stage 3 output of EML-DRC design

Table 2. Performance of existing methods in comparision to the proposed EML-DRC design

Method	Specificity	Accuracy
EML-DRC	98.34	98.38
RBM	95.5	89.47
K-NN and SVM	91	86.67
SVM and KNN	81.16	82
XG Boost Model	81.1	88.9

Table 3. Comparision of percentage of severity level using image processing and experts

S. no	Retinal fundus image	Percentage of severity level of diseased using image processing	Expertsvalidation percentage of severity level of diseased
1		0%	0 % Vision loss No DR
2		22%	Up to 25 % vision loss Mild DR
3		46 %	Up to 50 % vision loss Moderate DR
4		73%	50 TO 75 % loss in vision Severe DR
5		90 %	75 to 100 % loss in vision Proliferative DR
6		65%	50 TO 75 % Vision loss

4. CONCLUSION

The proposed EML-DRGC design has devised an effective diabetic retinopathy (DR) Diagnosis.The main objective of EML-DRGC design is to diagnose instinctively Diabetic Retinopathy with high specificity and high accuracy. Earlier research has implemented classification of DR stages using structural features, texture features, image features.In the proposed method novel efficient machine learning diabetic retinopathy grading classification (EML-DRGC) using multi SVM has been employed and additional action with this

accuracy and sensitivity have been improved.In addition to the structural features in the earlier research statistical GLCM and FAST corner detector features are included to strengthen the algorithm.Obtained results using the proposed EML-DRC Design highlights the outcomes in terms of various measures when compared to existing methods.In future EML-DRC Design will be realized into application of smart phone using IoT environment.

REFERENCES

[1] K. Xu, D. Feng, and H. Mi, "Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image," *Molecules*, vol. 22, no. 12, p. 2054, Nov. 2017, doi: 10.3390/molecules22122054.

[2] M. M. Abdelsalam, "Effective blood vessels reconstruction methodology for early detection and classification of diabetic retinopathy using OCTA images by artificial neural network," *Informatics in Medicine Unlocked*, vol. 20, p. 100390, 2020, doi: 10.1016/j.imu.2020.100390.

[3] N. Eladawi *et al*., "Early diabetic retinopathy diagnosis based on local retinal blood vessel analysis in optical coherence tomography angiography (OCTA) images," *Medical Physics*, vol. 45, no. 10, pp. 4582–4599, Oct. 2018, doi: 10.1002/mp.13142.

[4] T. R. Gadekallu *et al*., "Early detection of diabetic retinopathy using pca-firefly based deep learning model," *Electronics (Switzerland)*, vol. 9, no. 2, p. 274, Feb. 2020, doi: 10.3390/electronics9020274.

[5] M. D. Abràmoff and M. Niemeijer, "Mass screening of diabetic retinopathy using automated methods," in *Teleophthalmology in Preventive Medicine*, Berlin, Heidelberg: Springer, 2015, pp. 41–50.

[6] M. Bhaskaranand *et al*., "Automated diabetic retinopathy screening and monitoring using retinal fundus image analysis," *Journal of Diabetes Science and Technology*, vol. 10, no. 2, pp. 254–261, Mar. 2016, doi: 10.1177/1932296816628546.

[7] J. J, K. S. Umadevi, and R. J. Kannan, "Automated detection of microaneurysmsusing probabilistic cascaded neural network," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 11, no. 3, pp. 1083-1093, Sep. 2018, doi: 10.11591/ijeecs.v11.i3.pp1083-1093.

[8] W. Zhang *et al*., "Automated identification and grading system of diabetic retinopathy using deep neural networks," *Knowledge-Based Systems*, vol. 175, pp. 12–25, Jul. 2019, doi: 10.1016/j.knosys.2019.03.016.

[9] I. P. Okuwobi, Z. Ji, W. Fan, S. Yuan, L. Bekalo, and Q. Chen, "Automated quantification of hyperreflective Foci in SD-OCT with diabetic retinopathy," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 1125–1136, 2020, doi: 10.1109/JBHI.2019.2929842.

[10] S. Gayathri, A. K. Krishna, V. P. Gopi, and P. Palanisamy, "Automated binary and multiclass classification of diabetic retinopathy using haralick and multiresolution features," *IEEE Access*, vol. 8, pp. 57497–57504, 2020, doi: 10.1109/ACCESS.2020.2979753.

[11] M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C. M. Lim, E. Y. K. Ng, and A. Laude, "Computer-aided diagnosis of diabetic retinopathy: A review," *Computers in Biology and Medicine*, vol. 43, no. 12, pp. 2136–2155, Dec. 2013, doi: 10.1016/j.compbiomed.2013.10.007.

[12] A. P. Bhatkar and G. U. Kharat, "Detection of diabetic retinopathy in retinal images using MLP classifier," in *Proceedings - 2015 IEEE International Symposium on Nanoelectronic and Information Systems, iNIS 2015*, Dec. 2016, pp. 331–335, doi: 10.1109/iNIS.2015.30.

[13] M. D. Abràmoff *et al*., "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning," *Investigative Opthalmology & Visual Science*, vol. 57, no. 13, p. 5200, Oct. 2016, doi: 10.1167/iovs.16-19964.

[14] S. Kumar and B. Kumar, "Diabetic retinopathy detection by extracting area and number of microaneurysm from colour fundus image," in *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2018, pp. 359-364, doi: 10.1109/SPIN.2018.8474264.

[15] M. C. Boucher *et al*., "Diabetic retinopathy screening," *Canadian Journal of Ophthalmology*, vol. 44, no. 1, pp. 100–101, Feb. 2009, doi: 10.3129/i08-192.

[16] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, Jan. 1996, doi: 10.1016/0031-3203(95)00067-4.

[17] R. Haralick, K. Shanmugam, and D. Its'Hak, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, pp. 610–621, 1973, doi: 10.1109/TSMC.1973.4309314.

[18] A. Baraldi and F. Parmiggiani, "Investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 293–304, Mar. 1995, doi: 10.1109/36.377929.

[19] D. J. S. Raj, "Machine learning implementation in cognitive radio networks with game-theory technique," *IRO Journal on Sustainable Wireless Systems*, vol. 2, no. 2, pp. 68–75, May 2020, doi: 10.36548/jsws.2020.2.002.

[20] B. Antal and A. Hajdu, "An ensemble-based system for automatic screening of diabetic retinopathy," *Knowledge-Based Systems*, vol. 60, pp. 20–27, Apr. 2014, doi: 10.1016/j.knosys.2013.12.023.

[21] A. Mitra, "Sentiment analysis using machine learning approaches (Lexicon based on movie review dataset)," *Journal of Ubiquitous Computing and Communication Technologies*, vol. 2, no. 3, pp. 145–152, Sep. 2020, doi: 10.36548/jucct.2020.3.004.

[22] M. Chetoui, M. A. Akhloufi, and M. Kardouchi, "Diabetic retinopathy detection using machine learning and texture features," in *2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE)*, vol. 2018-May, pp. 1–4, doi: 10.1109/CCECE.2018.8447809.

[23] T. T. Ramanathan, M. J. Hossen, M. S. Sayeed, and J. E. Raja, "A deep learning approach based on stochastic gradient descent and least absolute shrinkage and selection operator for identifying diabetic retinopathy," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 25, no. 1, pp. 589–600, Jan. 2022, doi: 10.11591/ijeecs.v25.i1.pp589-600.

[24] L. K. Lok, V. Abdul Hameed, and M. Ehsan Rana, "Hybrid machine learning approach for anomaly detection," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 27, no. 2, pp. 1016-1024, Aug. 2022, doi: 10.11591/ijeecs.v27.i2.pp1016-1024.

[25] R. Mahakud, B. K. Pattanayak, and B. Pati, "Internet of things and multi-class deep feature-fusion based classification of tomato leaf disease," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 25, no. 2, pp. 995-1002, Feb. 2022, doi: 10.11591/ijeecs.v25.i2.pp995-1002.

[26] L. K. Pampana and M. S. Rayudu, "Quantification of retinal artery-vein ratio for vascular disease diagnosis using spatial U-Net," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 27, no. 3, pp. 1404-1411, Sep. 2022, doi: 10.11591/ijeecs.v27.i3.pp1404-1411.

[27] T. Kauppi *et al*., "The DIARETDB1 diabetic retinopathy database and evaluation protocol," in *Procedings of the British Machine Vision Conference 2007*, 2007, pp. 15.1-15.10, doi: 10.5244/C.21.15.
[28] M. Bader Alazzam, F. Alassery, and A. Almulihi, "Identification of diabetic retinopathy through machine learning," *Mobile Information Systems*, vol. 2021, pp. 1–8, Nov. 2021, doi: 10.1155/2021/1155116.
[30] Y. Zhao *et al*., "Using machine learning techniques to develop risk prediction models for the risk of incident diabetic retinopathy among patients with type 2 diabetes mellitus: a cohort study," *Frontiers in Endocrinology*, vol. 13, May 2022, doi: 10.3389/fendo.2022.876559.

BIOGRAPHIES OF AUTHORS

Srilaxmi Dasari is a Assistant Professorin the Department of Electronics and Instrumentation Engineering at VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad. She Has 15 years of teaching experience. She Holds a M. Techdegree in Electronics and Instrumentation Engineering with specialization in Electronics and Control. Her research areas are Image/signal processing, Artificial Intelligence, Medical image analysis and Computer Vision. She is pursuing Phd in the area of Biomedical ImageProcessing. Her research interests include image/signal processing, IoT, medical image analysis, and pattern recognition. She can be contacted at email: srilaxmi_d@vnrvjiet.in.

Boo Poonguzhali is a Assistant Professor in the Department of Electronics and Instrumentation Engineering at Annamalai University, Chidambaram. She Received her PhD from Annamalai Univeristy. She Holds a M. Techdegree in Process Control and Instrumentation. Her Specialization includes Vhdl, Process Control, Human Resource Management.She has 16 years of teaching experience including 5 years of Research.Her Research Interest includes process control instrumentation, machine learning, deep learning, biomedical imaging. She is a Life Member of Indian Society for Technical Education. She can be contacted at email: bookuzhali@gmail.com.

Manjula Sri Rayudu is a Professor in the Department of Electronics and Instrumentation Engineering at VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad. She Received her PhD from JNTUH, M.E from AU and BTech from JNTUK. She has 23 years of teaching experience including 10 years of Research. She has about 30 research articles published in various Journals and Conference Proceedings. She has three funding projects from AICTE & UGC and two Patents published under her Credit. She is a Senior Member IEEE, Member IEEE WIE AG, Member IEEE EMBS, Fellow of IETE, Life Member of Indian-Society for Technical Education, Life member of Instrumentation-Society of India (ISOI). She can be contacted at email: manjulasree_r@vnrvjiet.in.