Research Article

Identification of Diabetic Retinopathy Using Machine Learning

Dr. Satya Prakash Singh[a] Saumya Agrawal [b], Kuber Gupta[b], Purushottam Kaushik[b]

[sp.singh@bitmesra.ac.in] [saumyaagrawal490@gmail.com] [kubergupt80@gmail.com] [purshottamkaushik730@gmail.com]

[a]Assistant Professor, Computer Science & Engineering, Birla Institute of Technology, Mesra, Ranchi

[b]Student, B Tech. Information Technology, Birla Institute of Technology, Mesra

Abstract: Diagnosing diabetic retinopathy (DR) with colour fundus images is a difficult and time-consuming task due to a complex grading system and the demand for qualified doctors to determine the existence and importance of multiple microscopic characteristics. In this work, we propose a CNN approach for appropriately assessing DR severity from digital fundus images. We build a network with CNN architecture and data augmentations that can identify the intricate components necessary for the classification task, such as micro-aneurysms, exudate, and retinal hemorrhages, and then automatically offer a diagnosis without user input. We use a top-tier graphics processing unit (GPU) to train our network using the publicly available Kaggle dataset, and the results are excellent, especially for a challenging classification test. Our suggested CNN achieves a sensitivity of 95% and an accuracy of 75% on 5,000 validation images on the data set of 80,000 photos used.[1]

Key Words: Diagnosing diabetic retinopathy (DR), CNN approach, graphics processing unit (GPU), COCO, SVM.

1. Introduction

One of the main causes of blindness in the western world is diabetic retinopathy (DR). The number of persons with diabetes is expected to keep growing due to causes such as rising life expectancy, luxurious lifestyles, and other contributing variables. It has been demonstrated that routinely checking diabetic patients for DR is an essential and cost-effective part of their management. The timing and accuracy of this care have a critical role in the cost and efficacy of the treatment. This is a crucial step since, if diagnosed early enough, effective treatment for DR is available.

The weighting and placement of various features are necessary for classifying DR. This takes a lot of clinicians' time. Once educated, computers can make classifications significantly more quickly, improving the real-time classification assistance for clinicians. Computer imaging research has been actively investigating the effectiveness of automated grading for DR, with positive results. The detection of DR characteristics using automated techniques like support vector machines and k-NN classifiers has received a lot of attention. Most of these classification methods use a two-class system for DR or no DR.

A subset of deep learning known as convolutional neural networks (CNNs) has a remarkable track record for use in image processing and interpretation, including medical imaging. Network architectures created specifically to cope with picture data were regularly constructed with meaningful applications starting in the 1970s, and they outperformed other methods for difficult jobs like handwritten character recognition. Yet it wasn't until a number of advances in neural networks They were able to solve more challenging picture identification issues thanks to innovations like dropout, rectified linear units, and the associated rise in computing power provided by graphical processor units (GPUs). Currently, massive CNNs are used to successfully complete extremely challenging picture recognition tasks with a wide range of object classes. Several of today's cutting-edge image classification tasks, including the yearly ImageNet and COCO challenges, use CNNs.[2]

Within automated grading, and specifically CNNs, there are two basic problems. A suitable offset between sensitivity (patients accurately identified as having DR) and specificity (patients correctly identified as not having DR) is one goal. For national criteria, which is a five-class problem into normal, mild DR, moderate DR, and severe DR, this is significantly harder, and numerous DR classes. In neural networks, overfitting is a significant problem as well. The network overfits to the class that is most prevalent in the dataset when the dataset is skewed. Massive skewedness is common in large datasets. Less than 3% of the photographs in the dataset that we utilized were from the 4th and 5th grade, therefore modifications to our network were necessary to ensure that it could still learn the properties of these images.

In this study, we present a CNN method based on deep learning for the classification of DR in fundus images. As was mentioned previously, this is a medical imaging task with growing diagnostic value that has been the focus of numerous investigations in the past. From what we can tell, This is the first paper to discuss the CNN-based five-class classification of DR. The CNN is modified using a number of new techniques to fit our huge dataset. The performance and capacities of our network are then examined.

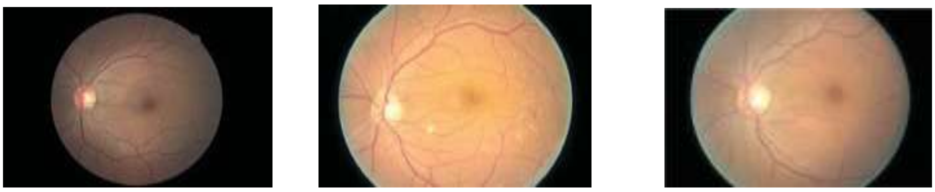
The rest of this essay is structured as follows. An overview of relevant research is provided in Section 2, the CNN's architecture and the training techniques are described in Section 3, the experiment results are shown in Section 4, and the paper is concluded with a discussion of the findings and recommendations for further research in Section 5.

2. Review of Related Studies

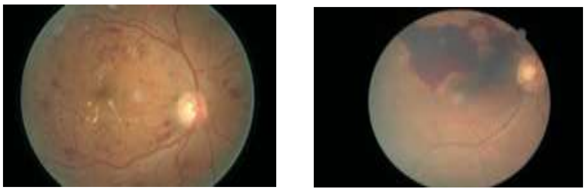
With hopeful outcomes, extensive study has been done on the techniques for a binary categorization of DR. Gardner et al. achieved sensitivity and specificity results for yes or no classification of DR of 88.4% and 83.5% using neural networks and pixel intensity measurements. They divided each image into patches using a small dataset of 200 photos, and before applying SVM, they required a physician to identify the patches according to certain criteria.

Moreover, DR has been classified into three categories using neural networks. Nayak et al combined textural metrics with features like the area of exudates and the amount of blood vessels. To categories pictures into normal, non-proliferative, and proliferative retinopathy, features are fed into a neural network. These features served as the input for the neural network's classification algorithm. By contrasting the detection results with expert ophthalmologists' ratings, the results were validated. They showed a 93% classification accuracy, a 90% sensitivity, and a 100% specificity. This was done on a dataset of 140 photos, and feature extraction, which can take some time, was needed on each image in both the training and testing phases.

Support vector machines have been employed in the great majority of studies on the five-class classification that have been conducted (SVMs). The five classes may be recognized automatically thanks to a method developed by Acharya et al. Using a higher-order spectrum approach, features are retrieved from the raw data, which represent the variation in the shapes and contours in the images, are supplied into the SVM classifier. [3]



(a) No DR (b) Mild DR (c) Moderate DR



(d) Severe DR (e) Proliferative DR

Fig 1: Stages of diabetic retinopathy (DR) with increasing severity

This SVM approach reported average accuracy, sensitivity, and specificity of 82%, 82%, and 88%, respectively. By calculating the areas of several features, including hemorrhages, micro-aneurysms, exudate, and blood vessels, Acharya et al. 19 have developed a five-class classification approach. Using image processing techniques, the features deemed to be the most significant, including blood vessels, micro-aneurysms, exudates, and hemorrhages, were retrieved from the raw pictures. After that, the SVM received these for classification. With the help of this approach, a sensitivity of 82%, specificity of 86%, and accuracy of 85.9% were attained. These approaches were developed on comparatively short datasets, and the loss in sensitivity and specificity was probably caused by the five class problem's complexity.

The automated diagnostic for DR created by Adarsh et al. included the detection of retinal blood vessels, exudate, micro-aneurysms, and textural

Fig 2: Network architecture

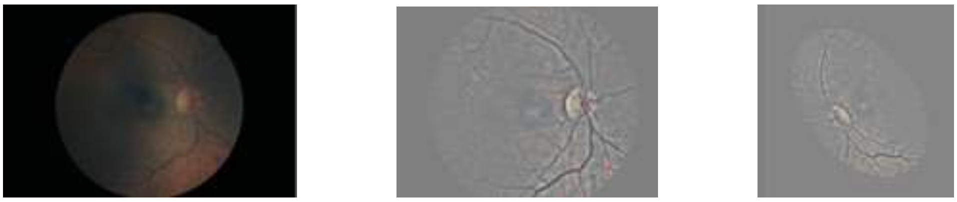
features. The feature vector for the multi-class SVM was built using the lesion area and texture features. Using the public 89 and 130

picture datasets, DIARETDB0 and DIARETDB1, respectively, this achieved accuracies of 96% and 94.6%.

Each of the first five class approaches needed to extract features from the photos before feeding them into an SVM classifier, and they were only successfully tested on tiny test sets of about 100 images. Compared to a CNN, these methods are less applicable in real-time. [3]

3. Method and Structure

We selected the structure of our neural network, which is shown in Fig 1, after evaluating the literature for different picture identification tasks. More convolution layers are thought to help the network learn more specific information. For example, whereas [12],[14]



(a) Original image (b) Preprocessed image (c) augmented image

Fig 3: Illustration of the preprocessing and augmentation processes

3.1. Preprocessing

Images from patients representing a wide range of ages, ethnicities, and lighting conditions were included in the dataset. As a result, needless variation that is unrelated to classification levels is created in the images' pixel intensity values. This was addressed by applying color normalization using the OpenCV (http://opencv.org/) module to the images. The outcome of this is depicted in Fig 3. (b). Furthermore having a high resolution, the photos took up a large amount of RAM. The dataset was downsized to 512x512 pixels so that the NVIDIA K40c could handle it while still preserving the complex features we wanted to discover.

3.2. Training

Prior to reaching a substantial level, the CNN underwent pre-training on 10,290 images. This was required in order to obtain a classification result pretty quickly without significantly increasing training time. The network was trained on all 78,000 training photos for an additional 20 epochs after 120 epochs of training on the initial images. With datasets like ours, where the bulk of the photos are categorized into one class—those displaying no indications of retinopathy—neural networks suffer from severe over-fitting. We created real-time class weights in the network to address this problem. The class-weights were modified for each batch loaded for back-propagation at a ratio corresponding to the number of photos in the training batch that were identified as showing no evidence of DR. This significantly decreased the probability of over-fitting to a certain class.

Using stochastic gradient descent and Nestrov momentum, the network was trained. The weights were stabilized using a low learning rate of 0.0001 over a period of five epochs. This was improved to 0.0003 for the lengthy 120 training epochs on the original 10,290 photos, increasing the model's accuracy to over 60%; this took about 350 hours of training. The network was then trained using a modest learning rate on the entire training set of images. The accuracy of the network reached over 70% after a few sizable epochs with the entire dataset. Each time training loss and accuracy reached a saturation point, the learning rate was subsequently reduced by a factor of 10.

3.3. Augmentation

The original pre-processed photos were used to train the network just once. The network's localisation capabilities were then improved during training by using real-time data augmentation. At each epoch, random rotations of 0–90 degrees, yes–or–no flips of the horizontal and vertical axes, and random shifts of the horizontal and vertical axes were applied to each image. Fig. 3 shows the outcome of an image augmentation (c).

4. Architecture of the CNN method for Segmentation of fundus images

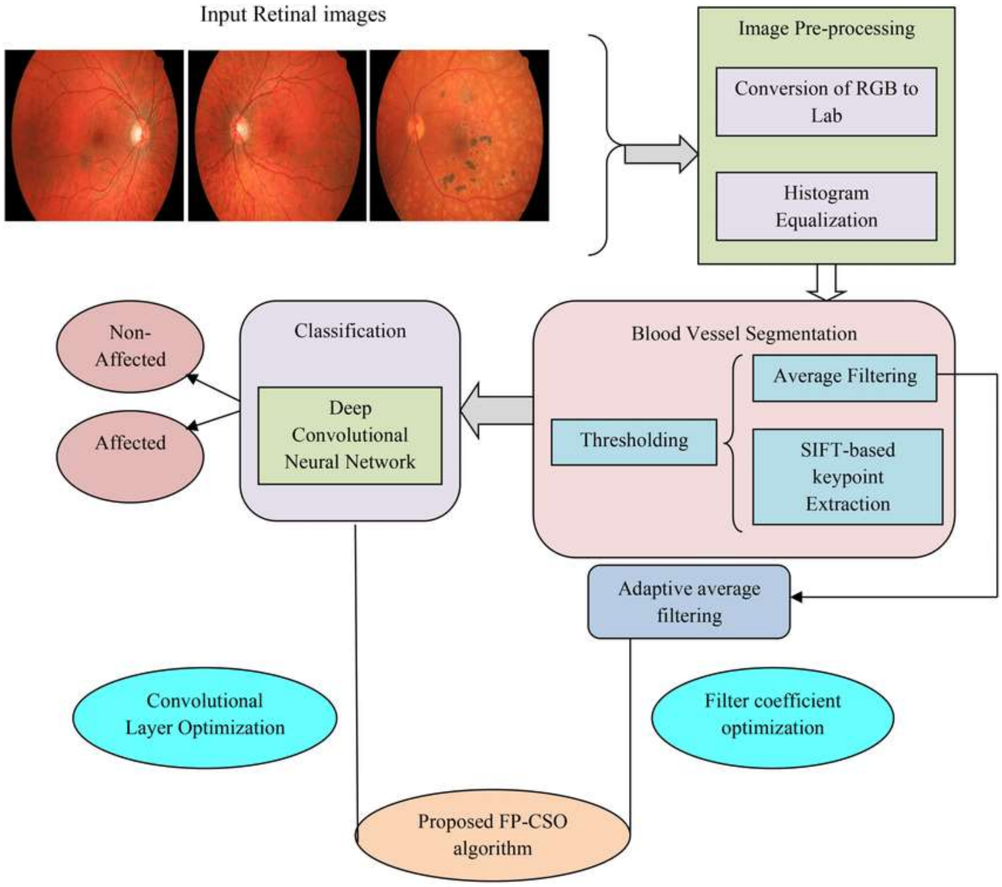


Fig 4: Architecture of the CNN Method for DR Detection

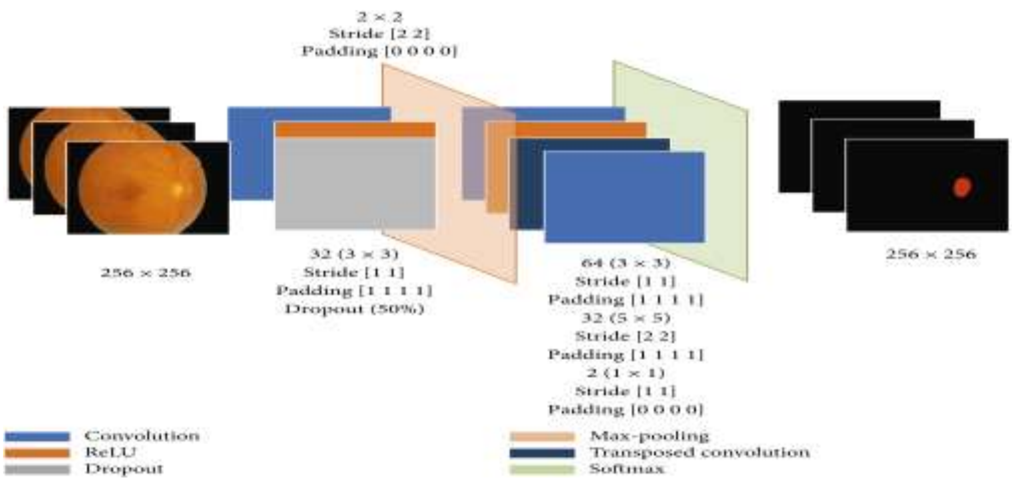


Fig 5: Segmentation of fundus images

5. Results

For validation reasons, 5,000 photos from the dataset were kept. It took 188 seconds to run the validation photos across the network. We define sensitivity as the number of patients correctly identified as having DR out of the genuine total amount of patients with DR and specificity as the number of patients correctly identified as not having DR out of the true total amount of patients with DR for this five-class problem. The number of patients with a correct classification is how we define accuracy. The final trained network has a final specificity of 95%, an accuracy of 75%, and a sensitivity of 30%. The network's classifications were defined numerically as follows: 0 - No DR Mild DR: 1 Moderate DR: 2 Three: Proliferative DR; four: Severe DR.[18]

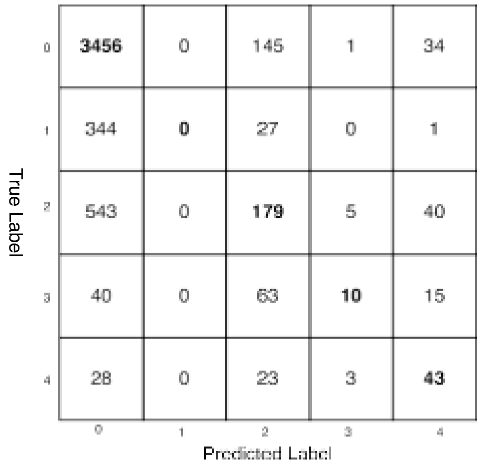


Fig 6: Confusion matrix of final classification results

Model Performance:

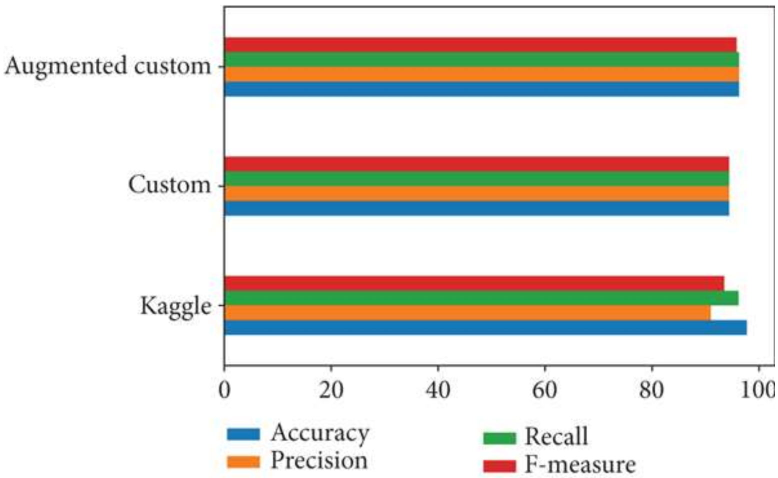


Fig 7: Model Performance Graph

Model Comparison: [19]

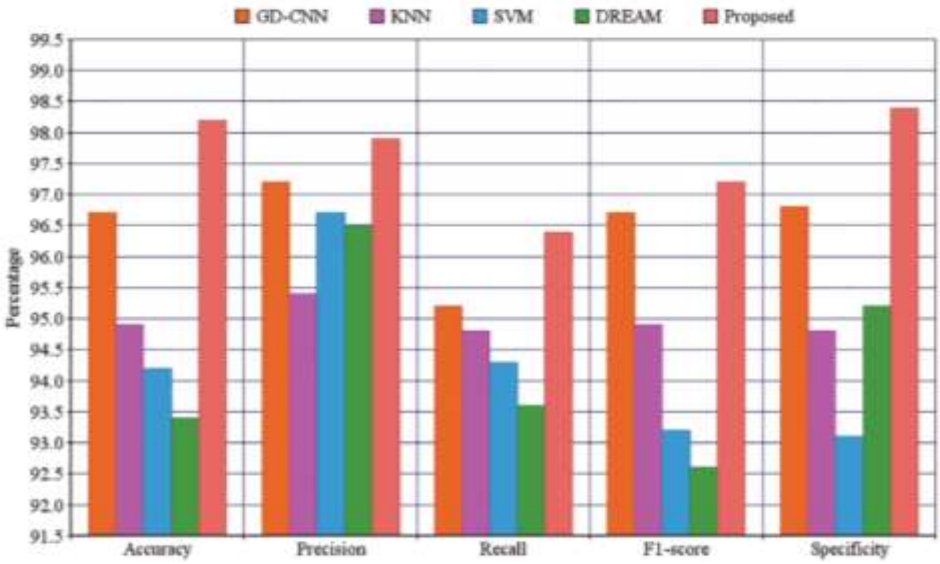


Fig 8: Performance comparisons of various ML algorithms which can be used for Diabetic Retinopathy Detection

6. Conclusion

Our research has demonstrated that the CNN method can be used to approach the five-class problem for the countrywide screening of DR. The majority of proliferative instances and those without DR were correctly classified by our network, which has showed encouraging signals of being able to learn the attributes needed to categories the fundus images. High specificity has been traded for poorer sensitivity, as in earlier research employing huge datasets. Our approach, which uses a substantially larger dataset and no feature-specific detection, achieves results that are comparable to those of these earlier approaches.

Our trained CNN can classify hundreds of images per minute, so using it in real-time whenever a new image is acquired could be useful. When a patient walks in for screening, in actuality, images are sent to doctors for grading and are not properly assessed. The trained CNN enables a quick diagnosis and a prompt response to a patient. The network also achieved these results by only using one image per eye.[8]

The picture of a healthy eye is quickly recognised by the network. This is most likely a result of the dataset's abundance of healthy eyes. Much less learning was required to classify the photographs at the extreme ends of the spectrum during training. There were issues while trying to get the network to distinguish between mild, moderate, and severe DR instances. The low sensitivity, which largely came from the mild and moderate classes, suggests that the network may have had trouble learning specific enough signals to identify some of the more complicated components of DR. A similar issue that was identified and validated by a clinician is that 10% or more of the images in our database are classified as ungradable by national UK criteria. These images were classified as a class because they contain a minimum quantity of DR. Given that the images were misclassified for both training and validation, our findings may have been seriously hampered. [18]

In the future, we intend to gather a dataset from actual UK screening setups that are significantly cleaner. As CNN continues to grow, deeper networks will be able to grasp the complex elements more easily than this network could. In terms of an established network topology, the results from our network are highly encouraging. Contrary to earlier techniques, nothing explicitly relating to the characteristics of our fundus images, such as vessels, exudate, etc., has been utilized. The CNN findings are outstanding as a result, but we have plans to tailor our network to this particular task in the future so that it can pick up on the subtler classification elements. We'll also compare these networks to five class SVM techniques that were developed using the same datasets.[13]

We have demonstrated that CNNs can be trained to recognize the characteristics of diabetic retinopathy in fundus images. As networks and databases continue to develop, CNNs will eventually be able to offer real-time classifications, which will be immensely helpful to DR physicians.

References:

1. Amalia, R.; Bustamam, A.; Sarwinda, D. Detection and description generation of diabetic retinopathy using convolutional neural network and long short-term memory. J. Phys. Conf. Ser. 2021, 1722, 12010.
2. Habibi, S., Ahmadi, M., and Alizadeh, S. (2015). Type 2 diabetes mellitus screening and risk factors using decision tree: results of data mining. Glob. J. Health Sci. 7, 304–310. doi: 10.5539/gjhs.v7n5p304
3. Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., and Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. Comput. Struct. Biotechnol. J. 15, 104–116. doi: 10.1016/j.csbj.2016.12.005
4. Jiang, H.; Yang, K.; Gao, M.; Zhang, D.; Ma, H.; Qian, W. An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification. In Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 2045–2048.
5. Assessing diabetic retinopathy from stereoscopic colour fundus photographs: an extension of the modified Airlie House classification.2018.
6. Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. Tang, “Predicting diabetes mellitus with machine learning techniques,” Frontiers in Genetics, vol. 9, 2018.
7. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
8. Sultana F, Sufian A, and Dutta P 2018 Advancements in Image Classification using Convolutional Neural Network IEEE 4th International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN 2018) 122-129
9. Priya H A G, Anitha J, and Rani J J N 2018 Computer Aided Diagnosis Methods for Classification of Diabetic Retinopathy Using Fundus Images IEEE International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET 2018) 1-4
10. Ioffe, S., Szegedy, C.. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* 2015;URL:arXiv:1502.03167.
11. He, K., Zhang, X., Ren, S., Sun, J.. Deep residual learning for image recognition. *arXiv* 2015;URL: arXiv:1512.03385.
12. Wang X, Lu Y, Wang Y, and Chen W B 2018 Diabetic Retinopathy Stage Classification using Convolutional Neural Networks IEEE 19th International Conference on Information Reuseand Integration for Data Science (IRI 2018) 465-471
13. Diabetic Retinopathy Classification Using CNN and Hybrid Deep Convolutional Neural Networks by Yasashvini R.,Vergin Raja Sarobin M.,Rukmani Panjanathan,Graceline Jasmine S. andJani Anbarasi L.ORCHID School of Computer Science and Engineering, Vellore Institute of Technology, Chennai 600127, India Symmetry 2022, 14(9), 1932; https://doi.org/10.3390/sym14091932
14. Nayak, J., Bhat, P.S., Acharya, R., Lim, C., Kagathi, M.. Automated identification of diabetic retinopathy stages using digital fundus images. *J Med Syst* 2008;**32**(2):107–115.
15. Acharya, R., Chua, C.K., Ng, E., Yu, W., Chee, C.. Application of higher order spectra for the identification of diabetes retinopathy stages. *J Med Syst* 2008;**32**(6):481–488.
16. Acharya, U., Lim, C., Ng, E., Chee, C., Tamura, T.. Computer-based detection of diabetes retinopathy stages using digital fundus images. *P I Mech Eng H* 2009;**223**(5):545–553.
17. Adarsh, P., Jeyakumari, D.. Multiclass svm-based automated diagnosis of diabetic retinopathy. In: *Communications and Signal Processing (ICCSP), 2013 International Conference on*. IEEE; 2013, p. 206–210.
18. Automated detection of diabetes using CNN and CNN-LSTM network and heart rate signals panelSwapna G, Soman Kp, Vinayakumar Centre for Computational Engineering and Networking (CEN), Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India Available online 8 June 2018, Version of Record 8 June 2018.
19. Intelligent Prediction Approach for Diabetic Retinopathy Using Deep Learning Based Convolutional Neural Networks Algorithm by Means of Retina Photographs G. Arun Sampaul Thomas1 , Y. Harold Robinson2 , E. Golden Julie3 , Vimal Shanmuganathan4 , Seungmin Rho5 and Yunyoung Nam6,*