Journal of Autonomous Intelligence (2023) Volume 6 Issue 2
doi: 10.32629/jai.v6i2.638

ORIGINAL RESEARCH ARTICLE

Classification and detection of diabetic retinopathy based on multi-scale shallow neural network

Mohamed Elageli M. Ghet[1,2,3,*], Omar Ismael Al-Sanjary[4,5], Ali Khatibi[6]

[1] *Department of Information Science & Computer, Management & Science University, Shah Alam 40100, Malaysia*
[2] *Higher Institute of Science and Technology, Qasr Ben Ghashir 21821, Libya*
[3] *Department of Computer Science, Al-Qasr International University, Qasr Ben Ghashir 21821, Libya*
[4] *Computer Center, University of Mosul, Mosul, Iraq*
[5] *Faculty of Information Science & Engineering, Management & Science University, Shah Alam 40100, Malaysia*
[6] *Post Graduate Centre, Management and Science University Shah Alam, Malaysia*
* **Corresponding author**: Mohamed Elageli M. Ghet, malajeele@hinstitute-bcv.edu.ly

ABSTRACT

The high-quality annotated training samples in medical image processing have limited the development of deep neural networks in their field. This paper designs and proposes an integrated method for classifying and detecting diabetic retinopathy based on a multi-scale shallow neural network. The method consists of multiple shallow neural network base learners, which extract pathological features under different receptive fields. The integrated learning strategy proposed is used to optimize the integration and finally realize the classification and detection of diabetic retinopathy. In addition, to verify the effectiveness of the method in this paper on a small sample data-set, based on the two-dimensional entropy of the image, multiple sub-datasets are constructed for verification. The results show that, compared with the existing methods, the integrated method for the classification and detection of diabetic retinopathy proposed in this paper has a good detection effect on a small sample data-set.

Keywords: image classification; diabetic retinopathy; multi-scale shallow neural network

ARTICLE INFO

Received: 17 May, 2023
Accepted: 29 May, 2023
Available online: 20 July, 2023

1. Introduction

Since 2000, approximately 151 million adults have had diabetes worldwide[1]. Today, 9.3% of adults between the ages of 20 and 80 have approximately 460 million people with diabetes[2]. In 2010, relevant organisations predicted that diabetes patients worldwide would reach 438 million by 2025. There are still four years, but it has exceeded the predicted value of 25 million people[3]. The most common complication in the patient population is diabetic retinopathy (DR), the leading cause of vision impairment and blindness in some adult patients[4]. Feature extraction is the key to image classification and detection. Furthermore, the grey histogram is a method of extracting image features in early research. It is widely used in image processing because of its simplicity and efficiency, Liu et al.[5] and Wang et al.[6] proposed the local binary mode texture feature extraction method, which is simple and easy to implement and has strong robustness to the change of grey level. Lu et al.[7] investigate reduction feature extraction's complexity and time consumption by extracting image pathological features and adding pathological location information to the feature space to improve performance. Thanh et al.[8] proposed a method to extract the shape feature of the lesion from the binary-segmented image according to the ABCD principle and track it with the original image to obtain the colour and texture features of the image. In clinical treatment, the diagnosis method for diabetic retinopathy is still by professional doctors observing the colour fundus images of the retina with the naked eye. So, the disease can be detected and treated early, and it can effectively help patients avoid blindness and reduce the blindness rate to 5%. Therefore, early detection and treatment of diseases are significant for diabetic patients and the world. This paper introduces an integrated method for the classification and detection of diabetic retinopathy proposed based on each of which is a shallow convolutional neural network. The detail proposes a wavelet transform-based fundus blood vessel segmentation method to achieve the feature extraction of fundus blood vessels at the same time. Combine the design of the shallow convolutional neural network model and propose a method that can be applied to small the classification and detection method of diabetic retinopathy of the sample data set.

2. Background

The current medical image processing field lacks high-quality annotated training samples, which leads to the fact that traditional machine learning algorithms and deep learning detection algorithms cannot achieve good detection results[9]. Ensemble learning combines some simple base learners for better results[10–13]. Usually, ensemble learning uses a single base learner and ensemble strategies such as the mean method and voting method to synthesise the output results of all base learners[12]. To sum up, the classification-detection ensemble method proposed in this paper consists of L-base learners to extract pathological features of different scales contained in retinal fundus images. In order to obtain a good classification and detection effect of diabetic retinopathy on a small sample data set, it is necessary to use an ensemble learning method to integrate the output results of each base learner to improve the detection accuracy. Currently, most of the existing ensemble learning methods are based on the average or voting method of the output results of all basic learners to optimize the detection effect. However, these ensemble methods are applied in multiple fields, such as machine learning and image classification. There seems to be little theoretical basis to explain their effectiveness. This paper introduces a simple and effective ensemble learning method, which optimizes the model detection effect by improving the performance of the strong base learner and reducing the performance of the weak base learner, which is called the performance ensemble strategy, which can be expressed as shown in Equations (1) and (2), respectively.

$$A_i = \frac{a_i}{\sum_{i=0}^{L-1} a_i} \tag{1}$$

$$C = \frac{\sum_{i=0}^{L-1}(A_i * Net_i(p_0, p_1, \ldots, p_{n-1}))}{L} \tag{2}$$

where, L is the number of basic learners, and a_i is the classification accuracy of the i basic learner, which means $\sum_{i=0}^{L-1} a_i$ is the sum of the classification accuracy of all basic learners.

Therefore, as shown in Equation (2), after determining the category of the input sample, A_i reflects the classification accuracy of the i basic learner and its influence on the total ensemble learning model; that is, A_i represents the i basic learner. The classification performance of each base learner accounts for the weight value of the performance of all base learners and satisfies the condition $0 \le i \le L-1$. $Net_i(p_0, p_1, \ldots, p_{n-1})$ represents the output result of the i basic learner when the sample is input to the ensemble learning model, where j_p represents the probability value that the input sample in the i basic learner belongs to the j class, and satisfies $(0 \le j \le n-1)$. Since diabetic retinopathy includes five lesion categories: no lesions, mild lesions, moderate lesions, severe lesions and proliferative lesions $(p_0, p_1, \ldots, p_{n-1})$, its output vector contains a total of $n = 5$ probability values. When an example is an input, C will decide the class it belongs to base on the outputs of all basic learners. The influence of the base learner on the classification outcomes of the model will be greater than that of other base learners, according to Equations (1) and (2), if the classification accuracy of the base learner is higher than that of other base learners. This is because the weight value of the base learner is larger than other base learners.

Furthermore, it is identified as a strong base learner. On the contrary, if its weight value A_i is small, it is regarded as a weak base learner, less influencing the final classification result. Therefore, based on the performance integration strategy proposed in this paper, the role played by strong base learners in the classification process will be enhanced, while the role played by weak base learners will be weakened. At present, there are still many challenges in the classification and detection of diabetic retinopathy in the field of medical image processing:

1) Manual detection method: The clinical treatment of diabetic retinopathy still relies on professional ophthalmologists for manual detection. This method not only consumes a lot of work-forces and material resources but also mainly relies on the prior knowledge of ophthalmologists. Knowledge and medical experience test results are subject to more significant subjective influence.

2) Machine learning detection method: The traditional detection methods based on machine learning algorithms are mainly based on clustering, supervised learning, statistical analysis and other methods. This type of detection method is not only susceptible to the influence of the algorithm itself but also requires manual feature labelling and feature selection.

3) Deep learning method: The deep learning-based diabetic retinopathy classification and detection system proposed by Google in 2019 have achieved good detection results[4], but its detection model not only has a large number of parameters and a complex structure but also requires a lot of the training samples are labelled and cannot be applied to small-sample datasets.

3. Proposed method

In order to offset this effect between features as much as possible, it is often necessary to normalize the data before model training to solve the comparability between features. After the original data is normalized, each feature vector is in the same order of magnitude, which is more conducive to speeding up the convergence speed of neural network training and improving the model training effect shown in Equation (3).

$$norm = \frac{x_i - min(x)}{max(x) - min(x)} \tag{3}$$

where, x_i represents the pixel value of the image, and $max(x)$ and $min(x)$ represent the maximum and minimum values of the image's pixels. Let the pixel value and the average gray value of the surrounding pixels be determined for each pixel in the image. The comprehensive qualities of the pixel, such as its gray level value and the distribution of neighboring pixels' grayscale values, will then be reflected in the binary group. Equation (3) illustrates how many times the binary group (i, j) appears in the image. That is, the frequency, M and N are the width and height of the image, respectively, then $P_{i,j}$ represents the probability of the occurrence of the dyad (i, j).

Therefore, the two-dimensional entropy of the image can be expressed as shown in Equations (4) and (5), respectively.

$$P_{i,j} = \frac{f(i,j)}{M \times N} \tag{4}$$

$$H = -\sum_{i=0,j=0}^{255} p_{ij} ln p_{ij} \tag{5}$$

The process of data pre-processing is mainly divided into four: the black border of the fundus image, meaningless images polluted by noise, and images with low light, before the classification and detection of diabetic retinopathy, as shown in **Figure 1**.

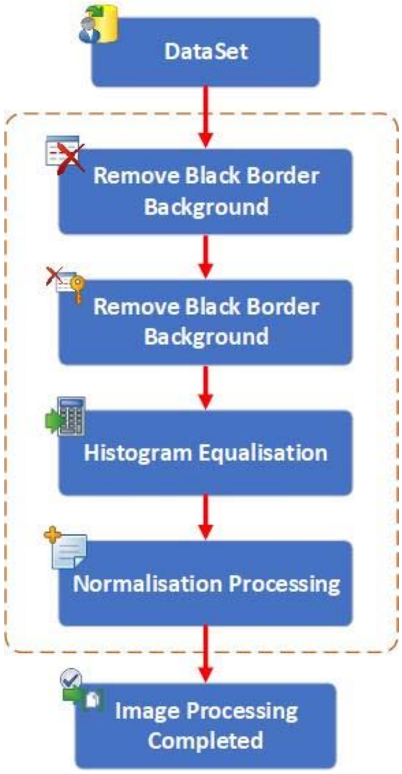


Figure 1. The pre-processing for a database.

Figure 2 displays the internal layers, pooling layers, and fully linked layers of three base learners as well as the general topology of a neural network model. The *L*-base learners have different convolution kernel sizes, which are used for the convolution layer to extract image features under different receptive fields. A pooling layer is added after the convolution layer to maintain data validity and alleviate overfitting.

Table 1 shows the pathological characteristics and data sample distribution characteristics of the 5 disease stages more clearly and intuitively. The classification criteria for lesions in diabetic retinopathy are as follows:

1) Net_0 is no lesions: Normal fundus, no diabetic retinopathy, such as no macula, no exudate, no retinal hemorrhage and other lesions.

2) Net_1 is mild lesions: Some pathological features such as yellow-white or grey spots in the retinal fundus image.

3) Net_2 is moderate lesions: Some pathological features such as yellow, red spots and exudates in retinal fundus images.

4) Net_3 is severe lesions: There are pathological features such as partial spots, exudates, and bleeding spots in the retinal fundus image.

5) Net_4 is proliferative lesion: Various pathological features such as spots, exudates, vascular or combined vitreous hemorrhages, and fundus micro-aneurysms in retinal fundus images.

4. Result and discussion

The diabetic retinopathy detection selects data-set Aptos from the Kaggle International Competition website for method validation and analysis. The data-set contains 35,126 colour RGB images of the retinal

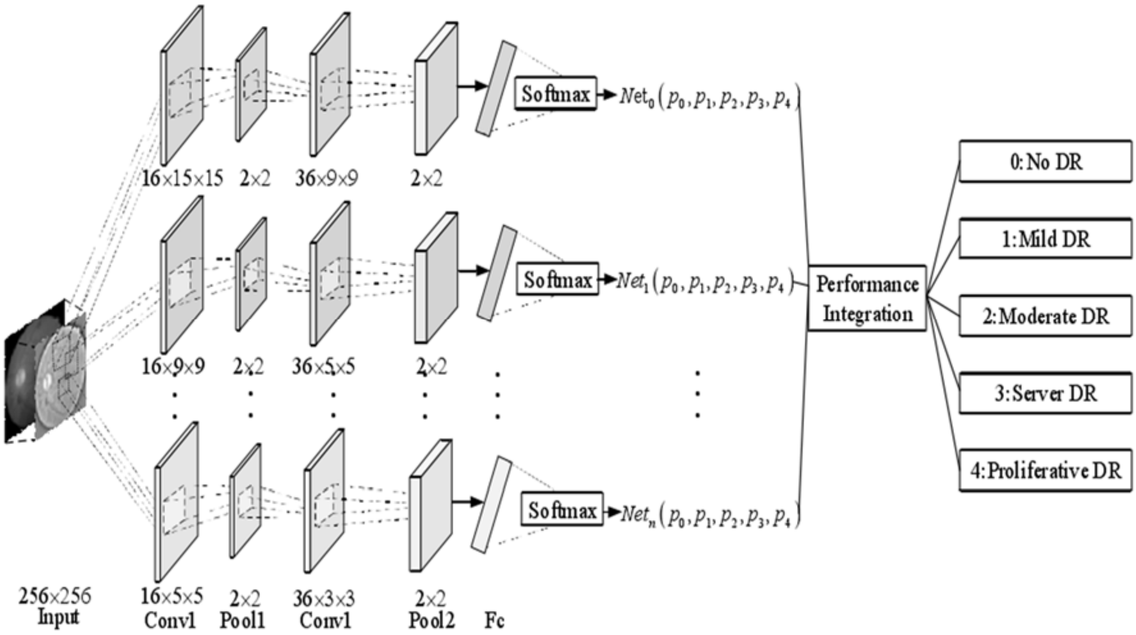


Figure 2. The proposed multi-scale shallow neural network model.

Table 1. Data set sample distribution table.

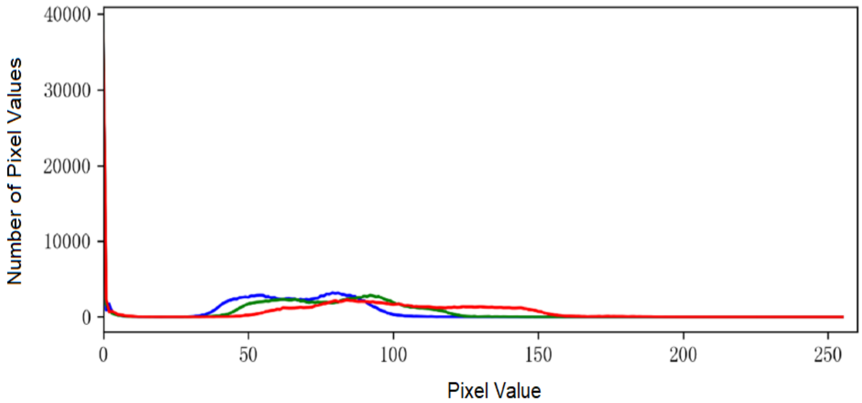
Category	Disease stage	Sample size
Net_0	No DR	25,810
Net_1	Mild DR	2443
Net_2	Moderate DR	5292
Net_3	Severe DR	873
Net_4	Proliferative DR	708

fundus and is divided into 5 disease stages according to the degree of diabetic retinopathy. **Figure 3** clearly shows the comparison of the histogram before and after equalization processing in the original image in **Figure 3(a)** cannot reflect the pathological characteristics of the retinal fundus very well due to insufficient lighting. Combined with the histogram data in **Figure 3(c)**. It can be seen that the pixels of the three channels of R, G, and B are mainly distributed in a small range between 50 and 150. After the histogram equalization pre-processing, the overall brightness of the fundus in **Figure 3(b)** is improved, the pixel distribution range of the three channels in **Figure 3(d)** is expanded to between 0 and 255, and the pixel distribution is relatively uniform. Finally, the channels can be combined to complete histogram equalization. In summary, the histogram equalization pre-processing operation enhances the dynamic range of the grayscale difference between pixels, which is more conducive to improving the classification and detection of diabetic retinopathy.

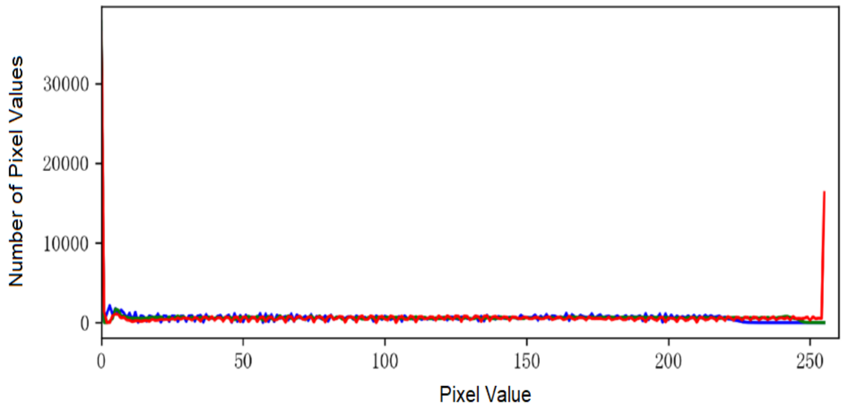
By way of illustration, using Net_0, the first convolution layer has 16 convolution kernels with a size of 13×13, and the second layer has 36 convolution kernels with a size of 8×8. The most recent Net_0 learner has fully connected 128-dimensional features. The training loss and accuracy of the five basic learners are displayed in **Figure 4** to make it easier to observe how they were trained using the same data-set and the multi-scale shallow neural network ensemble model. As the loss value falls during training, the accuracy curve of the base learner Net_0 shows an upward trend. The neural network converges, the accuracy increases to around 0.77, and the loss value decreases to about 0.75 when the number of iterations approaches 700. Similar tendencies may be seen in the loss of values and accuracy of the other basic learners. **Table 2** displays the categorization accuracy and model iterations of all basic learners. **Figure 4** and **Table 2** show that varied neural network model parameters do in fact influence the model's convergence speed and classification

(a) Original image

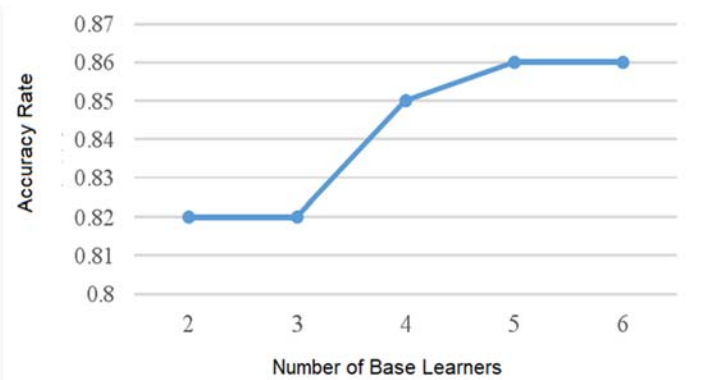
(b) Equalized image



(c) Original image histogram



(d) Image histogram after equalization processing

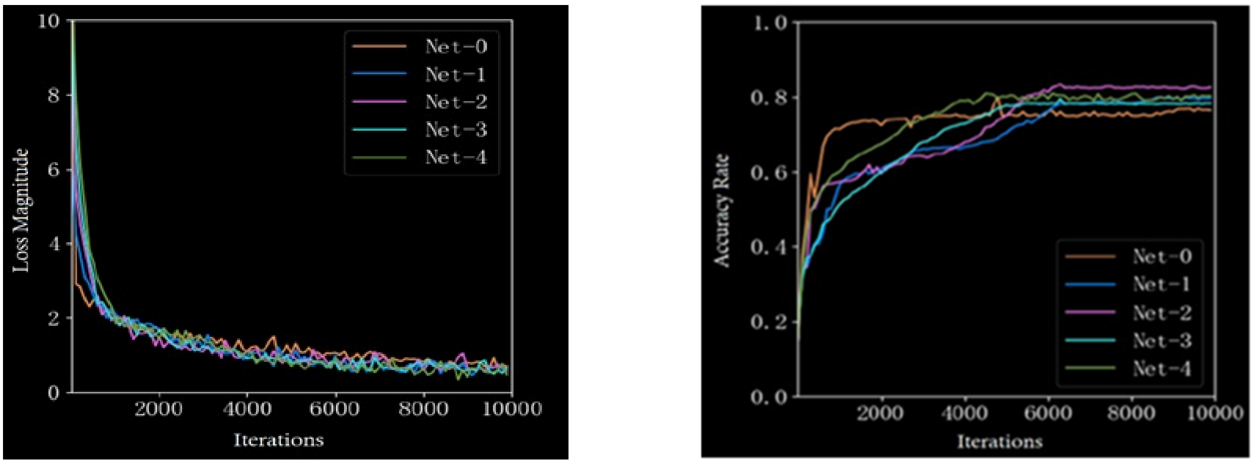


(e) The experiment—the number of base learners.

Figure 3. Histogram equalization processing.

Table 2. The number of iterations and accuracy of the base learner.

	Classification accuracy	Iterations number
Net_0	0.76	700
Net_1	0.79	800
Net_2	0.81	600
Net_3	0.77	700
Net_4	0.79	900



(a) Loss magnitude **(b)** Accuracy rate

Figure 4. The basic learner training.

Table 3. Basic learner network structure parameters.

	Convectional neural network	Multi-scale shallow neural network	Connected layer	Output result
Net_0	16 × 13 × 13	36 × 8 × 8	1 × 128	1 × 5
Net_1	16 × 7 × 7	36 × 5 × 5	1 × 128	1 × 5
Net_2	16 × 10 × 10	36 × 8 × 8	1 × 128	1 × 5
Net_3	16 × 6 × 5	36 × 6 × 5	1 × 128	1 × 5
Net_4	16 × 6 × 5	36 × 4 × 3	1 × 128	1 × 5

accuracy. This experimental finding demonstrates that it is possible to improve classification and detection of photos of diabetic retinopathy by applying the performance ensemble technique to optimize the output results of the multi-scale shallow neural network model.

In **Figure 4**, according to the previous exploratory experiments, when the number of basic learners of the ensemble model proposed in this paper increases from 2 to 5, its classification accuracy will increase from 82% to 86%. Nevertheless, when the number exceeds 5, the accuracy is no longer significantly improved. Therefore, while considering the efficiency of the ensemble model, this paper uses 5 basic learners to integrate the output results of the multi-scale shallow neural network ensemble model. The basic learners can be expressed as Net_i and $0 \leq i \leq 4$. The specific settings for each basic learner are listed in **Table 3**, including the convolution kernel size and the input and output of the fully connected layer.

In **Figure 5**, the classification accuracy of each base learner is lower than that of the 3 different ensemble models. Although Net_2 performs the best classification accuracy and convergence speed among all base learners, its results are still 1%–4% lower than the 3 ensemble models. According to the findings, the three ensemble methods' average classification accuracy is 4.6% greater than that of the five base learners. This serves as another evidence that the multi-scale shallow neural network ensemble model is a practical and efficient way to enhance the classification effect for medical images.

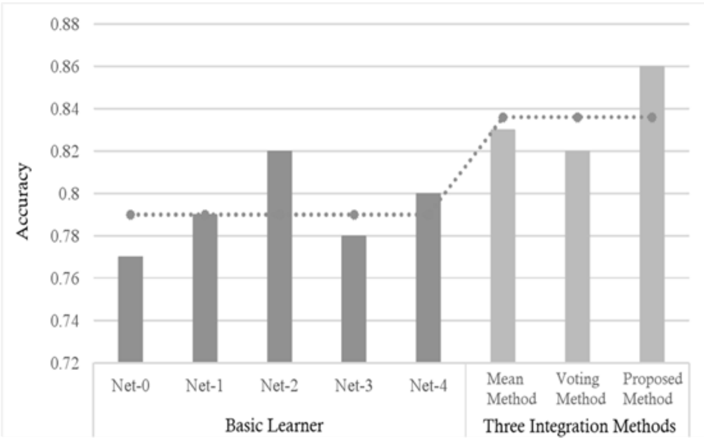


Figure 5. The accuracy comparison between the basic learner and different integration strategies.

5. Conclusion

This paper introduces a newly proposed method for classifying and detecting diabetic retinopathy images that can be applied to small sample datasets. The method consists of a multi-scale shallow neural network model proposed in this paper and the performance integration strategy. The multi-scale shallow neural network is used to extract pathological features under different receptive fields in retinal fundus images, and then the performance integration proposed in this paper is used. The strategy integrates and optimizes the output results of multiple base learners to obtain the final detection result. This method addresses the lack of high-quality annotated training samples in the current medical field. The experiment's result approved that the proposed method has a good detection effect on a small sample size of diabetic retinopathy image data set through multi-angle comparative analysis.

Author contributions

Conceptualization, MEMG and OIAS; methodology, MEMG; software, MEMG; validation, MEMG, OIAS and AK; formal analysis, MEMG; investigation, MEMG; resources, MEMG; data curation, MEMG; writing—original draft preparation, MEMG; writing—review and editing, OIAS; visualization, AK; supervision, OIAS, and AK; project administration, OIAS; funding acquisition, MEMG. All authors have read and agreed to the published version of the manuscript.

Acknowledgment

This study would not have been possible without the exceptional support of our supervisor. We would also like to thank the peer reviewers for their suggestions in preparing the manuscript. The generosity and expertise of one and all have improved this study in innumerable ways and saved me from many errors; those that inevitably remain are entirely responsible.

Conflict of interest

All authors have disclosed no conflict of interest. The funding bodies had no influence over the study's design, conduct, analysis, interpretation, article writing, or publication choice.

References

1. Sui M, Xue L, Ying X. Association of acupuncture treatment with mortality of type 2 diabetes in China: Evidence of a real-world study. *International Journal of Environmental Research and Public Health* 2020; 17(21): 7801. doi: 10.3390/ijerph17217801
2. Bilous R, Donnelly R, Idris I. *Handbook of Diabetes*. John Wiley & Sons; 2021.

3. World Health Organization. *WHO Global Report on Trends in Prevalence of Tobacco Smoking 2000–2025*, 2nd ed. World Health Organization; 2018.
4. Cavan D, Makaroff L, da Rocha Fernandes J, et al. The diabetic retinopathy barometer study: Global perspectives on access to and experiences of diabetic retinopathy screening and treatment. *Diabetes Research and Clinical Practice* 2017; 129: 16–24. doi: 10.1016/j.diabres.2017.03.023
5. Liu L, Fieguth P, Guo Y, et al. Local binary features for texture classification: Taxonomy and experimental study. *Pattern Recognition* 2017; 62: 135–160. doi: 10.1016/j.patcog.2016.08.032
6. Wang G, Sun J, Ma J, et al. Sentiment classification: The contribution of ensemble learning. *Decision Support Systems* 2014; 57: 77–93. doi: 10.1016/j.dss.2013.08.002
7. Lu S, Lu Z, Zhang Y. Pathological brain detection based on AlexNet and transfer learning. *Journal of Computational Science* 2019; 30: 41–47. doi: 10.1016/j.jocs.2018.11.008
8. Thanh DN, Prasath VBS, Hieu LM, et al. Melanoma skin cancer detection method based on adaptive principal curvature, colour normalisation and feature extraction with the ABCD rule. *Journal of Digital Imaging* 2019; 33(3): 574–585. doi: 10.1007/s10278-019-00316-x
9. Chen W, Yang B, Li J, et al. An approach to detecting diabetic retinopathy based on integrated shallow convolutional neural networks. *IEEE Access* 2020; 8: 178552–178562. doi: 10.1109/ACCESS.2020.3027794
10. Wang D, Wang X. The iterative convolution-thresholding method (ICTM) for image segmentation. *Pattern Recognition* 2020; 130: 108794. doi: 10.1016/j.patcog.2022.108794
11. Su H, Yu Y, Du Q, et al. Ensemble learning for hyperspectral image classification using tangent collaborative representation. *IEEE Transactions on Geoscience and Remote Sensing* 2020; 58(6): 3778–3790. doi: 10.1109/TGRS.2019.2957135
12. Nandhini N, Bhavani R. Feature extraction for diseased leaf image classification using machine learning. In: Proceedings of 2020 International Conference on Computer Communication and Informatics (ICCCI); 22–24 January 2020; Coimbatore, India. pp. 1–4.
13. Kumar A, Yaduvanshi RS. Quantum antenna operating at 430 to 750 THz band, inspired through human eye. *Journal of Information and Optimization Sciences* 2020; 41: 1365–1373. doi: 10.1080/02522667.2020.1809093