DE GRUYTER

Biomed. Eng.-Biomed. Tech. 2025; 70(3): 217–227

Bhuvaneswari Sekar* and Subashini Parthasarathy

Empirical analysis on retinal segmentation using PSO-based thresholding in diabetic retinopathy grading

Abstract

Objectives: Diabetic retinopathy (DR) is associated with long-term diabetes and is a leading cause of blindness if it is not diagnosed early. The rapid growth of deep learning eases the clinicians' DR diagnosing procedure. It automatically extracts the features and performs the grading. However, training the image toward the majority of background pixels can impact the accuracy and efficiency of grading tasks. This paper proposes an auto-thresholding algorithm that reduces the negative impact of considering the background pixels for feature extraction which highly affects the grading process.

Methods: The PSO-based thresholding algorithm for retinal segmentation is proposed in this paper, and its efficacy is evaluated against the Otsu, histogram-based sigma, and entropy algorithms. In addition, the importance of retinal segmentation is analyzed using Explainable AI (XAI) to understand how each feature impacts the model's performance. For evaluating the accuracy of the grading, ResNet50 was employed.

Results: The experiments were conducted using the IDRiD fundus dataset. Despite the limited data, the retinal segmentation approach provides significant accuracy than the non-segmented approach, with a substantial accuracy of 83.70 % on unseen data.

Conclusions: The result shows that the proposed PSO-based approach helps automatically determine the threshold value and improves the model's accuracy.

*Corresponding author: Bhuvaneswari Sekar, Department of Computer Science, Centre for Machine Learning and Intelligence (CMLI), Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India, E-mail: bhuvaneswari_cs@avinuty.ac.in. https://orcid.org/0000-0001-9777-1968
Subashini Parthasarathy, Department of Computer Science, Centre for Machine Learning and Intelligence (CMLI), Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, India, E-mail: subashini_cs@avinuty.ac.in. https://orcid.org/0000-0002-5786-4497

Introduction

According to the International Diabetes Federation (IDF), half a billion people worldwide have diabetes. According to estimations, 783.2 million adults worldwide will have diabetes in 2045 [1]. The primary cause of DR is a consequence of diabetes. DR leads to vision loss if it is not treated early. DR is primarily classified as either Non-Proliferative (NPDR or early DR) and Proliferative (PDR or late DR) [2]. The International Clinical Diabetic Retinopathy (ICDR) severity scale is a comprehensive tool for assessing the development of DR using fundus images. The inputted fundus images are classified into five categories according to the international protocol: Non-Diabetic Retinopathy (No DR (R0)), Mild Non-Proliferative Diabetic Retinopathy (NPDR (R1)), Moderate NPDR (R2), Severe NPDR (R3), and Proliferative DR (R4).

Based on severity level and referral requirements, the classification can also be divided into non-referable DR (No DR and Mild NPDR) and referable DR (Moderate NPDR, Severe NPDR, and PDR [3, 4]. Lesions evident on fundus imaging help distinguish between background DR and more advanced stages of the medical condition. The lesion feature generalization is difficult in Machine learning systems. Due to the growth in the high number of patients with DR, deep learning-based automatic screening of eye fundus images is the solution to screen a larger population of diabetes patients and it overcomes some limitations of automatic feature extraction and classification of the machine learning approach. However, the quality of the images affects the performance in the deep learning system [5]. To obtain a promising result, this study proposes Particle Swarm Optimization (PSO)-based auto thresholding algorithm for segment the retina from the background. There is limited work using retinal segmentation in DR Grading and Lesion Detection [6]. In addition, the importance of retinal segmentation is not shown in the prior work.

The primary objective of the paper is to propose PSO-based auto thresholding algorithms for retinal segmentation and to demonstrate the impact of the retinal segmentation

on the performance of the DR Grading System through the empirical assessment. The related objectives are described as follows:

- Applying the proposed PSO-based thresholding algorithms on fundus images to segment the retina.
- Comparing the effectiveness of proposed algorithms with the state-of-the-art Otsu, histogram-based sigma thresholding, and entropy-based thresholding algorithms.
- Implementing the pre-trained model ResNet50 to assess the impact of the non-segmented and retinal-segmented approaches in diabetic retinopathy grading.

The proposed work suggests PSO-based auto thresholding for retinal segmentation of fundus images. Then the retinal-segmented image is trained using transfer learning. Finally, the results are assessed to understand how the inclusion and exclusion of the segmentation impact the grading's performance. In this paper, the PSO-based thresholding algorithm is proposed to segment the foreground without manually setting the threshold, using entropy as an objective function. The PSO with local search and Entropy values are utilized to determine the positions of thresholds automatically.

The rest of the paper is structured as follows: In Section 2, relevant work on deep learning is reviewed. The dataset collection and proposed methods are covered in Section 3. Section 4 discusses the experimental outcomes. Section 5 presents the conclusion and future work.

Related work

An overview of the many methods utilized for DR Grading is given in this section. Early studies focused on improving the image quality using a standard approach of normalization, resizing, color conversion, and image augmentation.

The study [5] demonstrates the effectiveness of pre-processing in improving the performance of deep learning models for DR classification. The study, which concentrated on image augmentation and resizing to improve image quality, suggests that future research will include more preprocessing techniques. Feng et al. [7] combined Convolutional Neural Networks (CNN) and Graph Neural Networks (GNN) to extract the features and capture the complex relationships between data points, respectively. However, these models failed to focus on improving the quality of various dataset images.

To change the different sources images into a uniform format Gao et al. [8] utilized the resizing approach using

fixed dimensions, discarded the black borders by a selected threshold value, shape normalization to avoid the small notch on the edge of the image circle, normalized the image color to provide the more stable input range to the learning model, and image augmentation to overcome the limitations of data. Since the study utilized manual threshold setting to segment the retinal area. The study by Pao et al. [9] presents a Bi-channel CNN system by combining the features from both gray-level entropy images and green component entropy images to improve the detection of referable DR. To improve the image quality resizing, unsharp masking (UM), and augmentation is applied before computation without considering the additional preprocess approach.

Al-Turk et al. [10] described an approach based on feature-based grading to enhance transparency and interpretability in DR screening with the use of a simple pre-process approach of color averaging and augmentation. However, a lack of pre-processing techniques will affect the model performance and feature detection. In references [11–14], conventional pre-processing techniques are utilized to improve DR detection and grading. According to the study by Zhu et al. [15], the early studies focused on the primary preprocess of i) Image enhancement to overcome inhomogeneities of contrast using Contrast-limited Adaptive Histogram Equalization (CLAHE), AHE, and Modified CLAHE, ii) Denoising to remove the noise using the mean filter, gaussian filter, and wiener filter, iii) Normalization to reduce the feature bias, and iv) Augmentation to improve data size and diversity by applying geometric and color transformation.

Butt et al. [16] described an approach based on a hybrid method to extract features from the fundus images to improve the model performance in DR detection. In reference [17], custom training data were utilized to address varying image quality and class imbalance. The author introduced the pseudo-binary classification scheme to improve prediction reliability and provide more information, and also designed energy-efficient hardware, enabling deployment on edge devices. The study [18], evaluates the efficacy of two pre-trained models, Faster R-CNN and Mask R-CNN, for identifying hemorrhages and models were tested on an image database containing expert-labeled ground truth. Al-Hazaimeh et al. [19] proposed a novel method that combines artificial intelligence and image processing to enhance the identification of DR. The method aims to address issues with specificity, sensitivity, and accuracy in current techniques and successfully detects several forms of DR, including exudates, microaneurysms, and hemorrhages. The summary of the related work using deep learning approaches is shown in Table 1.

Table 1: Summary of related work using deep learning approaches.

References	Dataset	Method	Foreground retinal segmentation	Contributions
[5]	EyePACS	Image resize, and data augmentation	Not performed	To eliminate the overfitting, effective pre-processing and augmentation are carried out
[7]	APTOS2019, Messi-dor-2	Pre-process is not described	Not performed	Graph convolution network is introduced to learn the deep features of an image and its class information
[8]	Clinical dataset	Size, shape, and color normalization, data augmentation, and background cropping by manual threshold setting	Manual threshold setting is used	Proposed a novel dataset for clinical practice
[9]	Kaggle diabetic retinopathy	Data augmentation, resize, color component extraction	Not performed	The features of the gray-level entropy images and green component entropy images are combined using bi-channel CNN to enhance referable DR detection
[10]	Both clinical and public dataset	Color averaging, augmentation	Not described	An end-to-end feature-based DR grading and progression system is implemented
[11]	MESSIDOR, IDRiD	Size and color normalization, denoising, histogram equalization using CLAHE, augmentation	Not performed	Quadrant ensemble InceptionResnet V-2 framework is proposed for DR grading accuracy
[12]	IDRiD	Denoising, resize, augmentation, normalization	Fixed cropping is applied	A customized EfficientNet model is presented to utilize small datasets
[13]	DIARETDB0, DIA-RETDB1, Messidor, HEI-MED, ODIR	Resizing, normalization, augmentation	Not performed	The multi-classification deep learning model was proposed for diagnose different eye diseases
[14]	IDRiD, Messidor, DDR	Resize, augmentation	Not described	A cross-lesion attention network is proposed for DR grading
[16]	APTOS	Image resize, normalization	Not performed	The hybrid model of Google Net and ResNet18 is presented for binary classification and grading
[17]	EyePACS, DDR, APTOS	Filter out corrupted images and under-sampled the majority classes	Not described	Pseudo binary classification and energy efficient DR screening hardware device is designed for deployment
[18]	DIARETDB1	Methodology is not described	Not described	Detectron2 framework is utilized for providing the segmentation mask on hemorrhage
[19]	EyePACS	Denoised using non-linear wiener filter, OD is segmented using CHT for effective detection of exudates, the BCSPNFCM algorithm is introduced for blood vessel segmentation, and morphological operation is applied to make hemorrhage detection efficient	Not described	A new approach was introduced to increase the accuracy of DR detection by segmenting the features, and detection with the selection of lesions

The literature studies have proposed approaches to grading diabetic retinopathy (DR) in digital fundus images, but few considered retinal segmentation as another component that can enhance image quality. The current study addresses this gap by introducing a retinal segmentation technique based on Particle Swarm Optimisation (PSO) that aims to improve image quality and, consequently, the precision and robustness of automated DR grading systems. The general pipeline of the study is shown in Figure 1.

Materials and proposed methodology

Dataset collection and description

The current work utilized the Indian Diabetic Retinopathy Image Dataset (IDRiD) [20]. The IDRiD Dataset is a publicly accessible benchmark dataset. The fundus images were captured by a retinal specialist using a Kowa VX-10 alpha
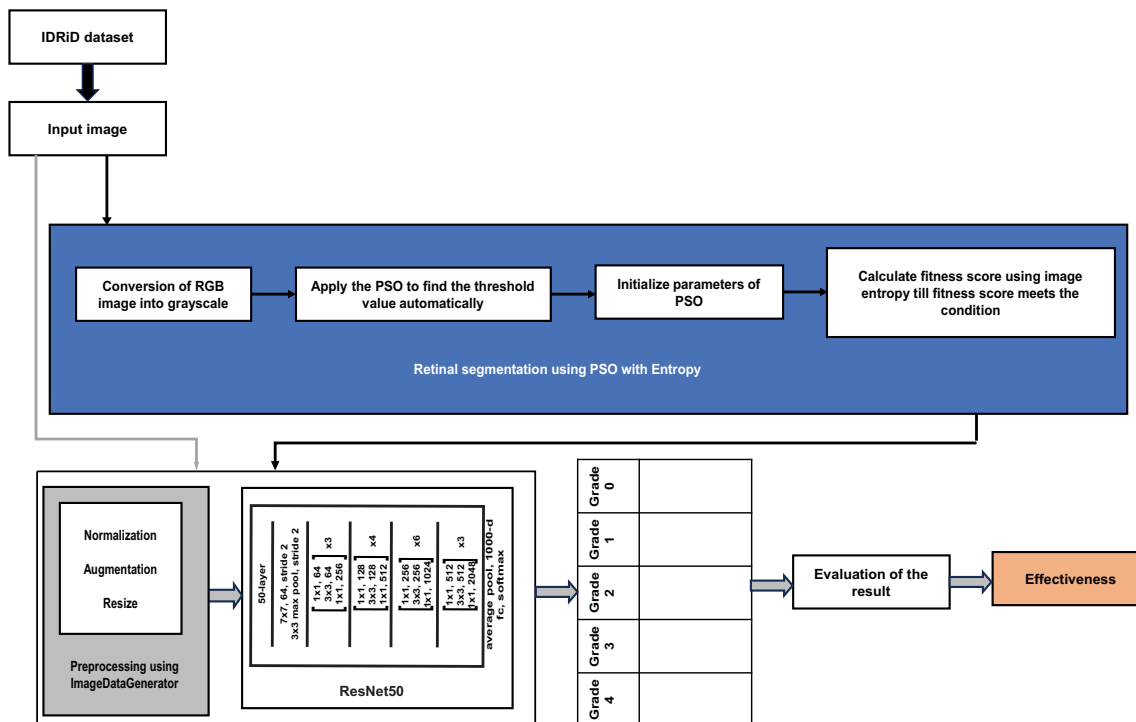
Figure 1: The general pipeline of the study.

digital fundus camera and all are centered near to the macula. The resolution of an image is $4,288 \times 2,848$ pixels. This dataset incorporates pixel-level annotated data and image-level grading data with binary mask ground-truth images to evaluate the performance of individual lesion segmentation techniques and grading classification.

Image pre-processing

The input data and the architecture of the networks play an important role in building an effective deep-leaning-based classification system. Pre-processing is an important step in improving the quality of the input image [21]. Improvement in preprocessing can help in eliminating artifacts [6]. This study has utilized normalization, augmentation, and resizing for preprocessing.

The image pixel values were originally in the range 0–255. These can lead to exploding or vanishing problems during the training process. Normalizing images to the range 0–1 helps to improve the image data, training stability, and performance of the model [22]. There are many normalization approaches for scaling the attribute values [23]. For this study, the rescale parameter from Keras ImageDataGenerator, a specific case of min-max normalization has used.

Image augmentation increases the size and diversity of the dataset artificially to improve the model's generalizability. The horizontal and vertical flipping is applied on the training image to augment the image using the Image-DataGenerator. The ImageDataGenerator avoids overfitting because it will not add the transformed images to the range of original images [24]. The image resize is used in pre-processing to transform different dimensional images to a consistent scale [8]. The image resize helps to minimize the running time and required memory of the model to load and train the images [25]. In this work, the fundus image is resized to downscale the image size by passing a fixed dimension value.

Proposed PSO-based retinal segmentation

Prior studies have made significant contributions to DR classification by enhancing image quality and facilitating feature analysis using various approaches including CLAHE, Color extraction, Feature extraction, Blood vessel segmentation, and Optic Disc segmentation. The artifacts may potentially cause the model to concentrate on inappropriate information. It is illustrated in the following Figure 2a using Explainable AI SHAP values tools. SHAP helps to understand
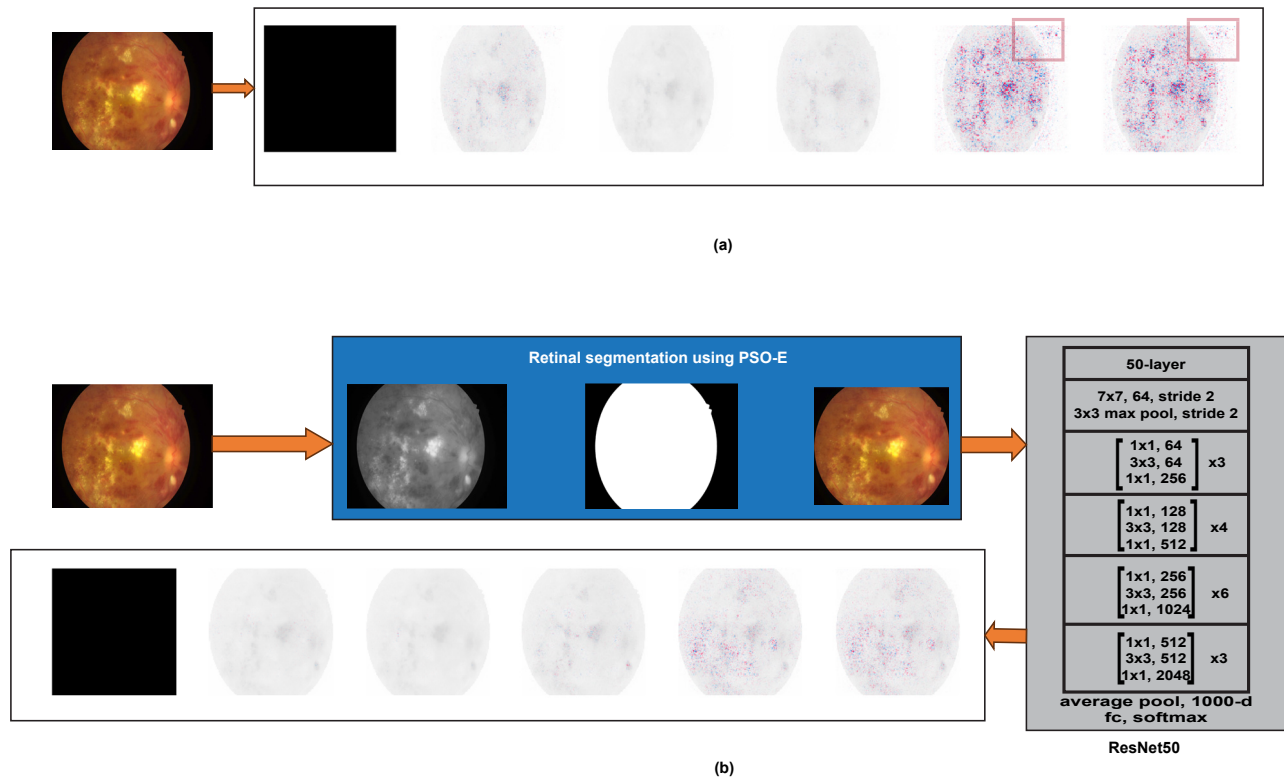
Figure 2: Result of SHAP interpretation on (a) non-segmented images and (b) segmented images using PSO-E.

how each feature impacts the model's prediction. SHAP gives an important value on each feature for a prediction that has been identified [26]. As emphasized in the figure, the model focuses on undesired black background information, impacting its accuracy and leading to misdiagnoses. This study suggests a retinal segmentation to segment the retinal area from the background area. This approach reduces background information to further enhance the model's accuracy and reduces the size of the image for more efficient processing. For foreground segmentation, thresholding is one of the most popular approaches for converting an image into a binary image. The optimal threshold value setting is the primary challenge in thresholding approaches, which is determined by the user. If the user is unable to choose an ideal threshold setting, it may cause to miss certain valuable areas of an image [27].

Thresholding techniques include local and global thresholding algorithms. Global thresholding uses a single threshold to divide the image into foreground and background pixels and this study focuses on global threshold approaches. Otsu's method is a popular state-of-the-art

global thresholding method, though it may struggle with certain images since it expects that the foreground and background have similar Gaussian distributions, which is a highly complex order that real-world images are unable to satisfy [28]. Otsu may not perform well with images that have more than two peaks in the histograms.

In recent studies, bio-inspired optimization algorithms have been utilized to tackle the problems of traditional optimization algorithms [29]. The bio-inspired PSO algorithm has the potential to produce good results while addressing the limitations of Otsu's technique, such as establishing an ideal threshold for accurate image segmentation and the images having a large intraclass variance difference and uneven illumination respectively. In 1990, Kennedy and Eberhart introduced the PSO [30]. PSO is a population and random-based approach inspired by the collective behavior of bird flocks or fish schools. The ideal strategy for the birds is to follow the one closest to the food. PSO treats each potential solution as a particle within the search space, much like a bird. Each particle has a fitness value, which is defined by a fitness function

that requires optimization, as well as a velocity that regulates its movement. Particles move through the problem space by seeking the particle in the best-known position [31].

For segmenting the retinal area from the unwanted background information, the PSO algorithm is used in this study due to its ease of implementation, low parameter needs, fast convergence, and scalability [32]. The purpose of using the PSO algorithm is to determine the threshold value automatically for the binary conversion of an image to segment the retinal region from images. In order to get the best threshold value, the entropy approach is used as an objective function with the PSO. The PSO with Entropy (PSO-E)-based optimization algorithm is implemented for auto threshold value determination and the detected threshold value is applied on the image to segment the retinal region. As shown in Figure 2b, the majority of background pixels are removed using the PSO-E algorithm, and the SHAP tool is applied to interpret the result. The SHAP shows the model is focused on the appropriate information after retinal segmentation. The general steps for the PSO-E algorithm for thresholding are as follows:

Step 1: Input the grayscale image, set the number of particles to 0, maximum iterations to 20, PSO parameters: inertia weight (w) to 0.5, cognitive and social components (c1, c2) are to1.5.

Step 2: Define the fitness function based on image entropy. The entropy of the binary image is used as the fitness to obtain the threshold value.

$$f(x_i) = -\sum_{j=1}^{N} p_j \log_2(p_j) \qquad (1)$$

where p_j is the probability of foreground pixel value in the binary image.

Step 3: Initialize the positions of particles with a predefined threshold range [10, 20], personal best positions are to be equal to the initial positions, global best position and fitness are to 0, all particle velocities are set to 0, and personal best fitness is to 0 for all particles.

Step 4: Iteration: begin;

Calculate the fitness of each particle.

Adjust the velocity and position of each particle.

$$v_i(t+1) = w.v_i(t) + c_1 \cdot r_1 \cdot (p_i - x_i) + c_2.r_2.(g - x_i) \qquad (2)$$

where

$v_i(t+1)$ is the velocity of the i – th particle.

$v_i(t)$ is the current velocity of the i – the particle.

g global best position found by any particle in the swarm.

p_i is the personal best position of the i – th particle.

Update personal and global best positions based on fitness values.

$$x_i(t+1) = x_i(t) + v_i(t+1) \qquad (3)$$

where

x_i is the current position of the i – th particle.

$x_i(t+1)$ is the new position of the i – the particle.

$x_i(t)$ is the current position the of i – th particle.

$v_i(t+1)$ is the new velocity of the i

– th particle calcuate from the first equation

Check the convergence criteria. Stop if the global fitness reaches the desired entropy of 1.0.

Step 5: Get the output of the binarized image using the global best threshold value.

Multi-class classification using ResNet50

The non-segmented and PSO-based retinal-segmented images are inputted separately into the ResNet50 transfer learning model [33] to evaluate the segmentation and grading performance. The Residual Neural Network is a Convolutional Neural Network (CNN) that integrates residual connections to preserve network connectivity and prevent gradient vanishing by adding the normalization layers. These connections work as shortcuts, allowing information to jump to certain levels and proceed directly to the output. This technique enables the network to learn residual functions and make modest parameter changes, resulting in faster convergence and better performance. ResNet enhances overall performance by prioritizing residual functions over complex input-output mappings.

The 50-layer ResNet (Residual Neural Network) architecture extracts a feature from the input image, and then the Global Average Pooling (GAP) [34] layer reduces the high dimensional feature maps from ResNet50 into lower-dimension vectors, and finally fully connected (dense) layers perform classification for grading the fundus images. This procedure aids in customizing the underlying model's higher-level features to the particular task at hand [35]. ResNet50 residual networks keep consistent computational complexity by doubling filters when halving feature map size and using the same filter count based on the output feature map size with bottleneck design to reduce the

number of parameters and matrix multiplications. Grading results are analyzed to provide insights into the significance of retinal segmentation.

Experimental results

Experimental dataset distribution

The experimental investigation of the proposed PSO-based retinal segmentation was executed on the images acquired from the IDRiD dataset. In total 516 images, for thresholding analysis, 516 images were employed, and for assessing the grading accuracy 330 images were allocated for training, 83 images for validation, and 103 images for testing. All computation, training, testing, and validations were performed in Google Colab environments using GPU runtime, equipped with an Intel i5 -1035G1 CPU @ 1.00 GHz processor, and 8 GB of physical memory hardware.

Performance metrics

To measure the performance of the retinal segmentation, kappa score, mean accuracy, multi-class confusion matrix, and SHAP XAI, are employed. In a multi-class confusion matrix rows reflect actual classes, and columns indicate expected classes. The kappa score is determined from the multi-class confusion matrix. It is more reliable than basic accuracy since it shows how much better the segmentation results are than probability. The segmentation model's mean accuracy provides an overview of how well it works across all classes, rather than concentrating on just one. The multi-class confusion matrix makes it possible to analyse errors for each class in greater detail and shows how frequently the model confuses one class with another. SHAP can explain which features are most crucial for predicting particular outcomes. The model's accuracy, efficiency, and reliability in retinal segmentation are evaluated using these metrics.

Experimental analysis of thresholding

Figure 3 presents the visualization results obtained from the state-of-the-art Otsu thresholding, histogram-based sigma thresholding, Entropy thresholding, and proposed PSO-E thresholding techniques.

The various thresholding algorithms such as Otsu's algorithm, the histogram-based sigma approach, the entropy-based thresholding algorithm, and the PSO optimization are effective for automatic thresholding methods.

However, Otsu, histogram-based sigma, and entropy methods lead to incorrect segmentation due to large intra-class variance differences and uneven illumination in images. In contrast, the proposed PSO algorithm, which uses entropy-based optimization, works well with both unimodal and multimodal histogram distributions. It also eliminates manual threshold setting, resulting in a more robust and efficient solution.

Experimental analysis of grading

The effectiveness of retinal segmentation is assessed using ResNet50 architecture. The various threshold techniques are applied to segment the retinal area. Once the retinal area was segmented, the segmented images were trained and validated using the classification model. The performance of different retinal segmentation methods was evaluated using training accuracy, validation accuracy and the mean kappa score, as shown in Table 2.

Based on the retinal segmentation table results, Otsu's method, while effective in training, suffered from over-fitting, the histogram-based approach had lower overall performance but a better balance between training and validation results, and the Entropy-based approach performed worse than Otsu or sigma-based approach, however, PSO-E-based thresholding indicates the highest agreement and showing good generalization on unseen data.

Ablation studies

To separately evaluate the impact of retinal segmentation, the deep learning model is trained using non-segmented images and PSO-based retinal-segmented images. Then the result is obtained separately using the ResNet-50 baseline model. The confusion matrix is a table that compares the grading performance of the two approaches, shown in Figure 4.

The figure shows retinal segmentation can significantly improve the accuracy of grading in deep learning. The findings suggest that retinal segmentation can help a deep learning model focus on the most important regions of an image, decreasing complexity and distractions. This can help the model learn crucial features and patterns for the task at hand. Table 3 (top) shows the class-wise performance.

From the above table, training with segmentation images yields superior results than non-segmentation images across all classes. Non-segmented images perform slightly better in validation in certain Class 0 and Class 3 cases. However, Segmentation approach provides a significant speed improvement for specific classes, such as Class 1
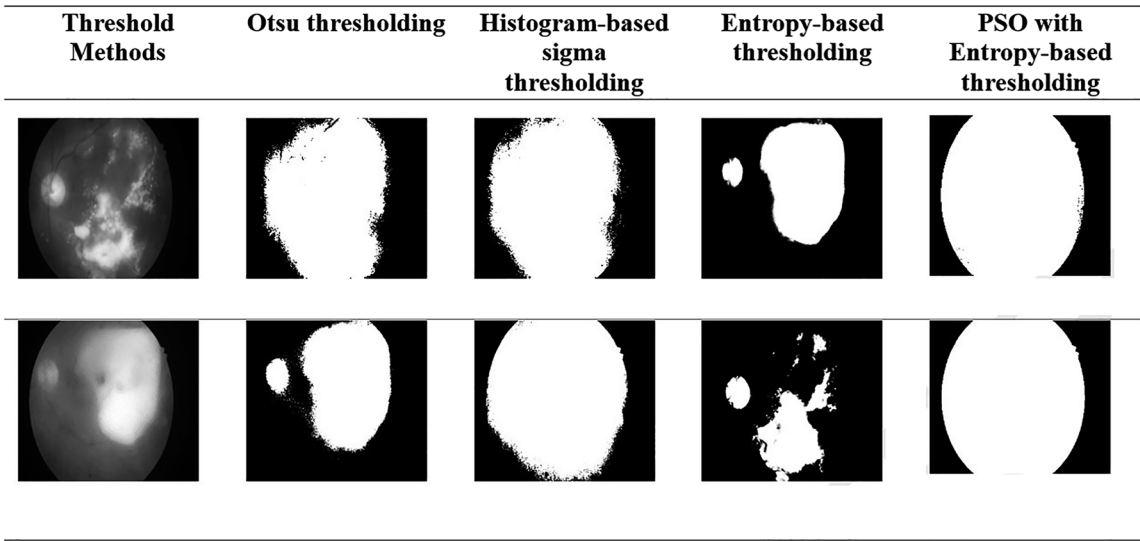
Threshold Methods	Otsu thresholding	Histogram-based sigma thresholding	Entropy-based thresholding	PSO with Entropy-based thresholding

Figure 3: Comparative performance of the thresholding with various methods.

Table 2: Comparative result on retinal segmentation.

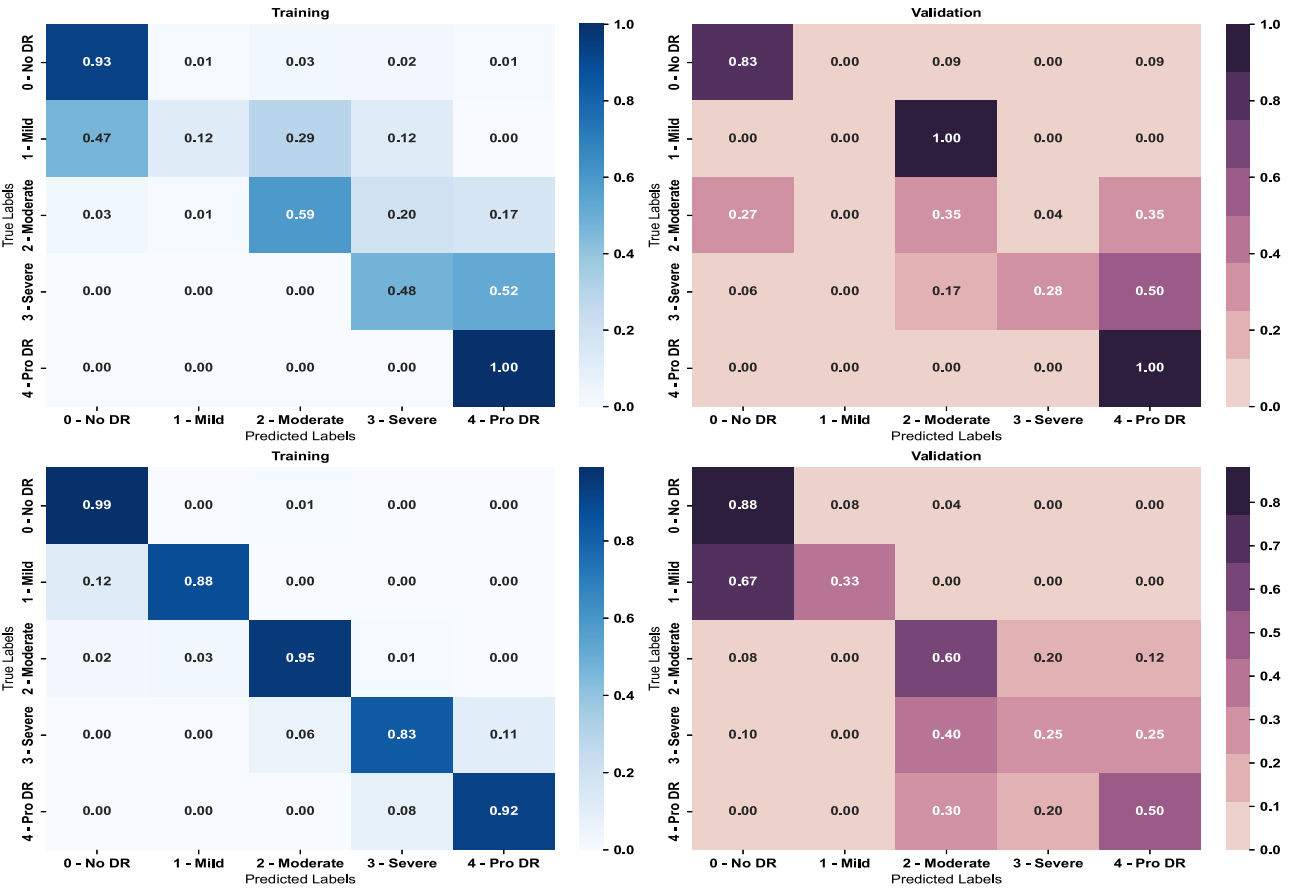| Methods	Training	Validation	Mean kappa score
Retinal-segmented images using Otsu	0.973	0.721	0.920
Retinal-segmented images using sigma-based	0.923	0.722	0.882
Retinal-segmented images using entropy-based	0.639	0.502	0.611
Retinal-segmented images using PSO with entropy	0.976	0.787	0.939

Figure 4: Result of multi-class confusion matrix on (top) non-segmented images (bottom) segmented images.

Table 3: Class-wise and overall accuracy results.

Approaches	Class-wise accuracy results											
	Class 0		Class 1		Class 2		Class 3		Class 4		Kappa score		Mean kappa score
	Training	Validation	Training	Validation	Training	Validation	Training	Validation	Training	Validation	Training	Validation	
Non-segmented images	0.89	0.96	0.80	0.79	0.83	0.61	0.85	0.74	0.91	0.85	93.6	69.5	88.6
Retinal-segmented images	**1.00**	**0.95**	**0.99**	**0.86**	**0.98**	**0.66**	**0.98**	0.72	**0.98**	**0.86**	**97.6**	**78.7**	**93.9**

Approaches	Overall accuracy		
	Training	Validation	Testing
Non-segmented images	0.879	0.817	0.817
Retinal-segmented images	**0.975**	**0.831**	**0.837**

Bolded values represent the highest values in each column, highlighting the best-performing metrics.

and Class 2. It is evident that performance is generally improved by retinal-segmented images. For training, validation, and overall performance of a task, the mean kappa score, a measure of agreement between expected and actual outcomes is assessed.

The model with the proposed retinal segmentation image accuracy is compared with the non-segmented image approach. The model trained using the proposed segmentation images gave an accuracy of 93.9 % using Kappa's score. The model with the non-segmented images obtained an accuracy of 88.6 %. Table 3 (bottom) indicates that the retinal-segmented approach performs much better than the non-segmented approach.

Compared to non-segmented images, the model performs much better when retinal-segmented images are used, especially when testing and validation stages are used to reflect real-world events. In order to improve generalization, the segmentation procedure aids the model in concentrating on important aspects inside the images. Figure 5 illustrates how test predictions increased for all classes when employing retinal-segmented images; nevertheless, the non-segmented approach was unable to meet this accuracy, especially for Grade 1, which is indicated by a red bounding box.

Finally, the experimental results show that the majority of background pixels affect the DR grading. Therefore, retinal segmentation using a PSO-based thresholding approach is proposed to segment the foreground retinal region from the background. Hence, the retinal segmented images not only reduce the complexity but also obtain a modest improvement in the performance of the model and grading accuracy.

Conclusions

Recently, deep learning has been utilized in medical images. It can uncover complex patterns and segment medical images efficiently. It saves time and cost on feature engineering in medical categorization to identify medical pictures. It has the greatest accuracy among all image prediction algorithms. However, early processing is important to achieve the best result in deep learning. In this paper, PSO-based auto thresholding proposed to segment the retina from the background, and the impact of the retinal segmentation in DR grading is analyzed. The proposed PSO-thresholding approach does not require any threshold value setting. The algorithm automatically determines the threshold value to segment the foreground retinal area and background unwanted information. The entropy method is used with PSO to determine the threshold value for segmentation. To
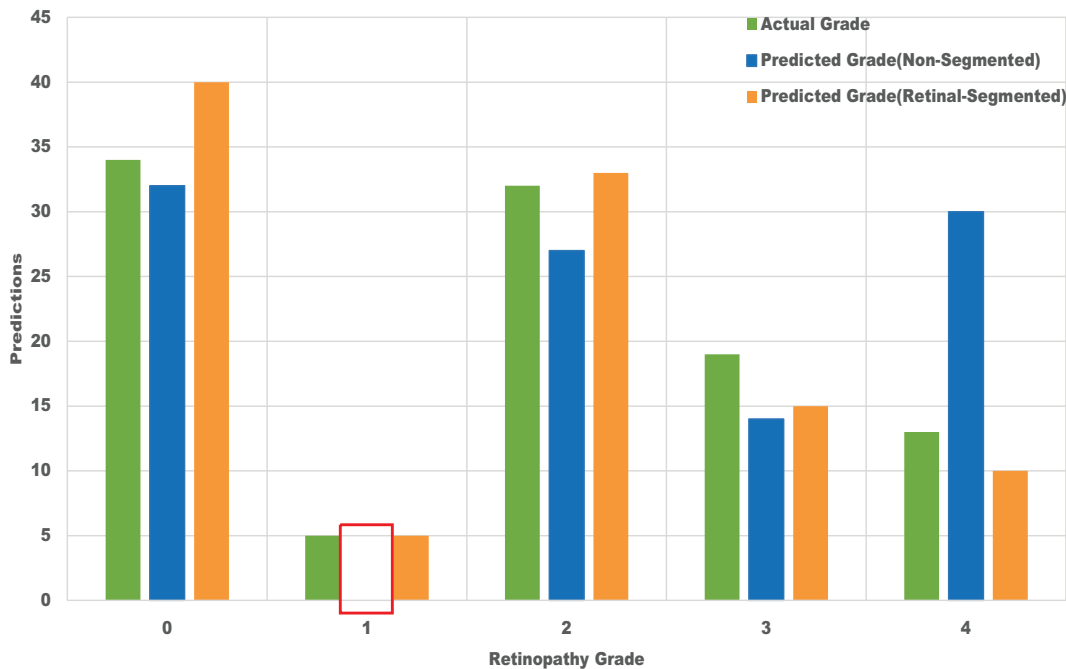
Figure 5: Prediction result on the test dataset.

evaluate the ability of the proposed PSO segmentation, the image using PSO thresholding and state-of-the-art thresholding is inputted and tested using the ResNet50 architecture. The confusion matrix, kappa's scores, and various performance metrics are considered for evaluation. The proposed approach improves the model's accuracy. Both approaches present a promising result, with the retinal segmentation approach yielding a higher precision than the non-segmented approach. To obtain better outcomes, future research work will concentrate on the development of the detection of optic disc and multi-level lesions of the DR signs with the largest number of images.

Research ethics: As we used a public dataset, not applicable.
Informed consent: Not applicable.
Author contributions: The authors has accepted responsibility for the entire content of this manuscript and approved its submission.
Conflict of interests: The authors state no conflict of interest.
Data availability: Publicly available datasets have been used in this study.

Declaration of AI and AI-assisted technologies in the writing process: During the preparation of this work, the authors utilized QuillBot to paraphrase and Grammarly to grammar. After using these tools, the authors reviewed and edited the content as needed and took full responsibility for the final publication.

References

1. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF Diabetes Atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract 2022;183:109119.
2. Raudonis V, Kairys A, Verkauskiene R, Sokolovska J, Petrovski G, Balciuniene VJ, et al. Automatic detection of microaneurysms in fundus images using an ensemble-based segmentation method. Sensors 2023;23:3431.
3. Raman R, Ramasamy K, Rajalakshmi R, Sivaprasad S, Natarajan S. Diabetic retinopathy screening guidelines in India: All India Ophthalmological Society diabetic retinopathy task force and Vitreoretinal Society of India consensus statement. Indian J Ophthalmol 2021;69:678–88.
4. UK Government. Your guide to diabetic retinopathy [Internet]; 2022. Available from: https://www.gov.uk/government/publications/diabetic-retinopathy-description-in-brief/your-guide-to-diabetic-retinopathy.
5. Incir R, Bozkurt F. A study on effective data preprocessing and augmentation method in diabetic retinopathy classification using pre-trained deep learning approaches. Multimed Tools Appl 2024;83:12185–208.

6. Jabbar A, Liaqat HB, Akram A, Sana MU, Azpíroz ID, Diez IDLT, et al. A lesion-based diabetic retinopathy detection through hybrid deep learning model. IEEE Access 2024;12:40019–36.

7. Feng M, Wang J, Wen K, Sun J. Grading of diabetic retinopathy images based on graph neural network. IEEE Access 2023;11:98391–401.

8. Gao Z, Li J, Guo J, Chen Y, Yi Z, Zhong J. Diagnosis of diabetic retinopathy using deep neural networks. IEEE Access 2018;7:3360–70.

9. Pao SI, Lin HZ, Chien KH, Tai MC, Chen JT, Lin GM, et al. Detection of diabetic retinopathy using bichannel convolutional neural network. J Ophthalmol 2020;2020:9139713.

10. Al-Turk L, Wawrzynski J, Wang S, Krause P, Saleh GM, Alsawadi H, et al. Automated feature-based grading and progression analysis of diabetic retinopathy. Eye 2022;36:524–32.

11. Bhardwaj C, Jain S, Sood M. Deep learning–based diabetic retinopathy severity grading system employing quadrant ensemble model. J Digit Imaging 2021;34:440–57.

12. Abdelmaksoud E, Barakat S, Elmogy M. Diabetic retinopathy grading system based on transfer learning. Int J Adv Comput Res 2021;11:1–12.

13. Albelaihi A, Ibrahim DM. DeepDiabetic: an identification system of diabetic eye diseases using deep neural networks. IEEE Access 2024;12:10769–89.

14. Liu X, Chi W. A cross-lesion attention network for accurate diabetic retinopathy grading with fundus images. IEEE Trans Instrum Meas 2023;72:1–12.

15. Zhu S, Xiong C, Zhong Q, Yao Y. Diabetic retinopathy classification with deep learning via fundus images: a short survey. IEEE Access 2024;12:20540–58.

16. Butt MM, Iskandar DNFA, Abdelhamid SE, Latif G, Alghazo R. Diabetic retinopathy detection from fundus images of the eye using hybrid deep learning features. Diagnostics 2022;12:1607.

17. Diware S, Chilakala K, Joshi RV, Hamdioui S, Bishnoi R. Reliable and energy-efficient diabetic retinopathy screening using memristor-based neural networks. IEEE Access 2024;12:47469–82.

18. Chincholi F, Koestler H. Detectron2 for lesion detection in diabetic retinopathy. Algorithms 2023;16:147.

19. Al-Hazaimeh OM, Abu-Ein AA, Tahat NM, Al-Smadi MA, Al-Nawashi MM. Combining artificial intelligence and image processing for diagnosing diabetic retinopathy in retinal fundus images. Int J Online Biomed Eng 2022;18:131–51.

20. Porwal P, Pachade S, Kamble R, Kokare M, Deshmukh G, Sahasrabuddhe V, et al. Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research. Data 2018;3:25.

21. Asia AO, Zhu CZ, Althubiti SA, Al-Alimi D, Xiao YL, Ouyang PB, et al. Detection of diabetic retinopathy in retinal fundus images using CNN classification models. Electronics (Basel) 2022;11:2740.

22. Gavali P, Banu JS. Deep convolutional neural network for image classification on CUDA platform. In: Deep learning and parallel computing environment for bioengineering systems. Cambridge, MA: Academia Press; 2019:99–122 pp.

23. Patro SG, Sahu KK. Normalization: a preprocessing stage. Int Adv Res J Sci Eng Technol 2015;2:20–2.

24. Sarki R, Ahmed K, Wang H, Zhang Y, Ma J, Wang K. Image preprocessing in classification and identification of diabetic eye diseases. Data Sci Eng 2021;6:455–71.

25. Abd Aziz N, Sulaiman MAH, Mohd Yassin AI, Megat Ali MSA, Abu Hassan H, M Shafie S, et al. Preprocessing of fundus images for detection of diabetic retinopathy. J Electric Electron Syst Res 2021;19:149–56.

26. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 2017;30:4768–77.

27. Rahkar Farshi T, Demirci R. Multilevel image thresholding with multimodal optimization. Multimed Tools Appl 2021;80:15273–89.

28. Zhou D, Xia Z. An improved Otsu threshold segmentation algorithm. J China Univ Metrol 2016;27:319–23.

29. Darwish A. Bio-inspired computing: algorithms review, deep analysis, and the scope of applications. Future Comput Inform J 2018;3:231–46.

30. Kennedy J, Eberhart R. Particle swarm optimization. Proceedings of ICNN'95. In: International Conference on Neural Networks. Perth, WA, Australia: IEEE; 1995, 4:1942–8 pp.

31. Bhandari S, Rambola R, Kumari R. Swarm intelligence and evolutionary algorithms for diabetic retinopathy detection. In: Swarm intelligence and evolutionary algorithms in healthcare and drug development. New York: Chapman and Hall/CRC; 2019:65–92 pp.

32. Shanthi MB, Meenakshi DK, Ramesh PK. Particle swarm optimization. In: Advances in swarm intelligence for optimizing problems in computer science. New York: Chapman and Hall/CRC; 2018:115–44 pp.

33. Bichri H, Chergui A, Hain M. Image classification with transfer learning using a custom dataset: comparative study. Procedia Comput Sci 2023; 220:48–54.

34. Lin M, Chen Q, Yan S. Network in Network. In: 2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, 14–16 April; 2014. http://arxiv.org/abs/1312.4400

35. TensorFlow. Transfer learning and fine-tuning; 2024. Available from: https://www.tensorflow.org/tutorials/images/transfer_learning.