Mathematics and Computer Science
2025; 10(4): 60-69
http://www.sciencepublishinggroup.com/j/mcs
doi: 10.11648/j.mcs.20251004.11

Segmented Retinal Fundus Image Classification Approach with Multiple Instance Learning

Mahbub Hasan*

Department of Computer Science and Engineering, Southeast University, Dhaka, Bangladesh

Email address:
mahbub.hasan@seu.edu.bd (Mahbub Hasan)
*Corresponding author

To cite this article:
Mahbub Hasan. (2025). Segmented Retinal Fundus Image Classification Approach with Multiple Instance Learning. *Mathematics and Computer Science, 10*(4), 60-69. https://doi.org/10.11648/j.mcs.20251004.11

Received: 1 August 2025; **Accepted:** 18 August 2025; **Published:** 25 September 2025

Abstract: Diabetic Retinopathy (DR) is a leading cause of vision loss, and its early detection is essential for preventing unchanging damage. This study proposes an efficient classification framework using segmented retinal fundus images from the MESSIDOR dataset, focusing on vessel-level segmentation to enhance disease-specific feature extraction and minimize extraneous background information. The pipeline includes preprocessing steps such as image resizing, green channel extraction, binary masking, and Contrast Limited Adaptive Histogram Equalization (CLAHE), followed by segmentation of both thick and thin vessels using traditional image processing techniques. Segmented features are classified within a Multiple Instance Learning (MIL) framework, evaluating MIL with Lagrangian Relaxation (MIL-RL), Semiproximal SVM (mi-SPSVM), Kernel-based Semiproximal SVM (mi-KSPSVM), and conventional SVM (linear and RBF kernels). Performance is evaluated using Testing Correctness (TC), Sensitivity, Specificity, F1-score, and computational time under 5-fold, 10-fold, and Leave-One-Out (LOO) cross validation (CV). Results show that preprocessing significantly improves accuracy across all models, with mi-SPSVM under LOO-CV achieving the best at 68.00% in TC and a 68.32% in F1-score. At the same time, mi-KSPSVM achieved the highest sensitivity of 84.00%. The findings reveal that a segmentation-first approach in the classification task can enhance both diagnostic accuracy and computational efficiency, making it suitable for real-time and resource-limited environments. This approach offers a worthwhile, scalable solution for automated DR screening, with future work aimed at integrating deep learning-based segmentation to further improve precision.

Keywords: Machine Learning, Optimization, Multiple Instance Learning, Support Vector Machine, Image Segmentation, Retinal Fundus Image

1. Introduction

Diabetes is a group of disease that often happen when blood sugar is increased [1]. Currently, there is an approximation of 598 million adults living with this disease. The hormone responsible for maintaining blood glucose levels. The condition disrupts normal circulation within the retina's delicate, light-sensitive blood vessels, leading to reduced or blocked blood flow. As a result, DR can cause significant visual problems, ranging from mild impairment to complete blindness. Therefore, early detection through routine eye examinations is crucial for individuals with diabetes, given their elevated risk of developing this condition [2]. AI-based treatment has emerged, resulting in a significant impact on the medical field. Only clinical trials are inadequate in early disease identification. Machine Learning (ML) and Deep Learning (DL) approaches significantly influence this sector in solving problems with practical solutions. This work emphasizes the use of fundus retinal medical images to aid in the identification of diabetes. We achieved an incredible milestone with the help of Multiple Instance Learning (MIL) and its various optimized algorithms. The progression of diabetes and its effects on the retina undergo several changes, each stage having distinct implications for vision and eye health.

1. *Early Stages*: In this stage, the tiny blood vessels within the retina may become damaged and cause

microaneurysms to develop. These microaneurysms can leak fluid into the retina, causing macular edema: a condition in which the center of the retina (the macula) [3] swells, thereby affecting central vision.

2. *Intermediate Stages*: As diabetes progresses, abnormal blood vessels grow due to a huge level of sugar in the blood. Therefore, a gel-like fluid flows at the back of the eye. This abnormal bleeding led to permanent damage to vision; in extreme cases, it can lead to retinal detachment.
3. *Advanced Stages*: This stage is diabetic macular edema (DME), a complication in which the macula becomes thickened and distorted due to fluid buildup. This causes vision loss in affected individuals, often affecting central vision.

Fundus imaging reveals several important biomarkers, including the optic disc, blood vessels, and lesions such as microaneurysms, hemorrhages, and hard and soft exudates. It serves as a valuable diagnostic tool for various ocular conditions, including glaucoma, diabetic retinopathy, age-related macular degeneration, retinopathy of prematurity, and diabetic macular edema [4]. Each of those biomarkers is visualized based on the type of disease. For example, glaucoma results from damage to the optic nerve, and segmenting the optic disc allows for more precise observation of anatomical changes associated with this condition. Figure 1 illustrates the possible biomarkers of a fundus image.

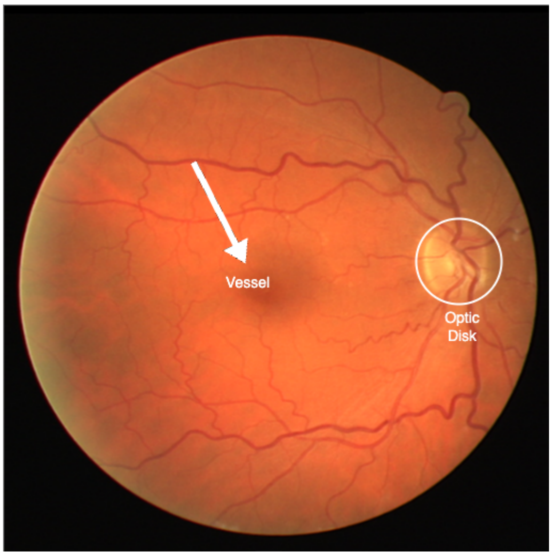

Figure 1. *A fundus image illustrating important biomarkers.*

Additionally, we place great importance on image segmentation in this research. Because image segmentation has a significant impact on achieving accurate disease classification. Most researchers are concentrating entirely on this sector, as the effect of segmented images on a specific area becomes more precise, clear, and focusable.

The rest of the paper is organized as follows: In Section II, we present several background analyses of our research. In Section III, we provide the design methodology of our proposed algorithm. In Section IV, we present the results and discussion. Section V contains the concluding remarks.

2. Literature Review

In recent years, biomedical imaging has gained considerable attention from researchers, with the human eye being a focal point of many studies. A wide array of techniques has been developed to detect key retinal features, such as blood vessels, optic discs, and important lesions. These approaches span supervised and unsupervised learning, multi-scale analysis, vessel tracking, mathematical morphology, thresholding, model-based methods, and other general strategies. Among those approaches, one of the most well-supervised approaches, known as MIL, is emerging in this field.

Andrews et al [5] first introduce MIL as a maximum-margin SVM and mixed-integer quadratic programs for the classification task. Based on that, Astorino et al. [6] proposed Lagrangian relaxation technique that was applied to the SVM model for binary MIL classification problem. Later on, [1, 7] extended this approach to color image analysis and preliminary studies for detecting DR from retinal fundus images.

Avolio and Fuduli developed an optimized MIL and proposed a semiproximal SVM (SP-SVM) method that combined with Kernel [8, 9]. Those approaches achieve high testing correctness on the benchmark dataset. To validate Lagrangian relaxation combined with block coordinate descent (BCD) and produced optimal solutions a binary MIL used by Astorino et al. [10]. However, further improvement using Multi-instance Nonparallel Tube Learning (MINTL) by Xiao et al. [11] incorporates large-margin tube embedded boundary information to improve generalization capability.

On the other hand, Zumpano et al. [12] used MIL to classify chest X-ray images and demonstrated MIL's potential in handling noisy clinical data. Similarly, Avolio et al. [2] applied MIL to classify retinal fundus images of healthy individuals and those with severe DR. These studies together demonstrate the increasing use of MIL for healthcare AI.

In the context of retinal imaging, DL approaches have significantly advanced for automatic diagnosis. Jeong et al. [13] focused on a review of DL methods for filtering diabetic retinopathy, glaucoma, and age-related macular degeneration (AMD). Li et al. [14] augmented this scope by outlining 143 publications and 33 publicly available datasets, showing a hierarchical summary of lesion segmentation, disease classification, and synthetic image generation. These studies provide a basis for choosing and assessing models and datasets in retinal analysis.

Sarki et al. [15] proposed a framework that incorporates classical image processing with a custom Convolutional Neural Network (CNN) to classify DED. Their model yielded high specificity and accuracy. For optic disk and cup segmentation, Edupuganti et al. [16] presented a Fully Convolutional Network (FCN) that improves cup and disk ratio computation for glaucoma diagnosis. Ahmad Fadzil et al. [17] proposed a vessel extraction and reconstruction algorithm that used adaptive histogram equalization and Gaussian-based region growing. Their method achieved excellent performance on the DRIVE dataset. Vessel enhancement work by Frangi et al. [18] and Mahapatra et al. [19] enhanced the multiscale Frangi filter using a modified particle swarm optimization and embedding it with fuzzy clustering, resulting in superior segmentation outcomes.

Lara and Urcid Serrano [20] employed Fourier and cosine transforms to segment both blood vessels and exudates using frequency-domain filtering and morphological operations. Huang and Kumar [21] developed a complete model that included segmentation and measurement of vessel attributes like length, width, and tortuosity – crucial parameters for neonatal diagnosis. Dataset availability and benchmarking are important to the rise of computerized diagnostic models. Decencière et al. [22] discussed the impact of the Messidor dataset on the research community, emphasizing its importance. Odeh et al. [23] used the Messidor dataset to test ensemble machine learning models that combined feature selection and classification. Mateen et al. [3] presented a broad review of DR detection systems, categorizing methods by dataset, performance, and algorithm type.

Beyond this theoretical understanding and benchmarking, Carbonneau et al. [24] offered a systematic and in-depth review of MIL by categorizing its problem characteristics into four dimensions: bag composition, ambiguity in instance labels, data distribution, and task-specific formulations. This theoretical framework has been used to contextualize the practical deployment of MIL in complex domains.

3. Methodology

Image segmentation serves as a core process in computer vision, aiming to divide an image into distinct regions based on specific areas of interest. Its primary objective is to group pixels or regions that exhibit similar characteristics, allowing for clearer analysis and understanding of the image content. This research solely focuses on the image segmentation task using fundus retinal images. Once the segmentation task is completed, image feature extraction is applied, followed by optimization algorithms to classify images of healthy and weak patients. Overall pipeline of this research is illustrated in the following Figure 2.

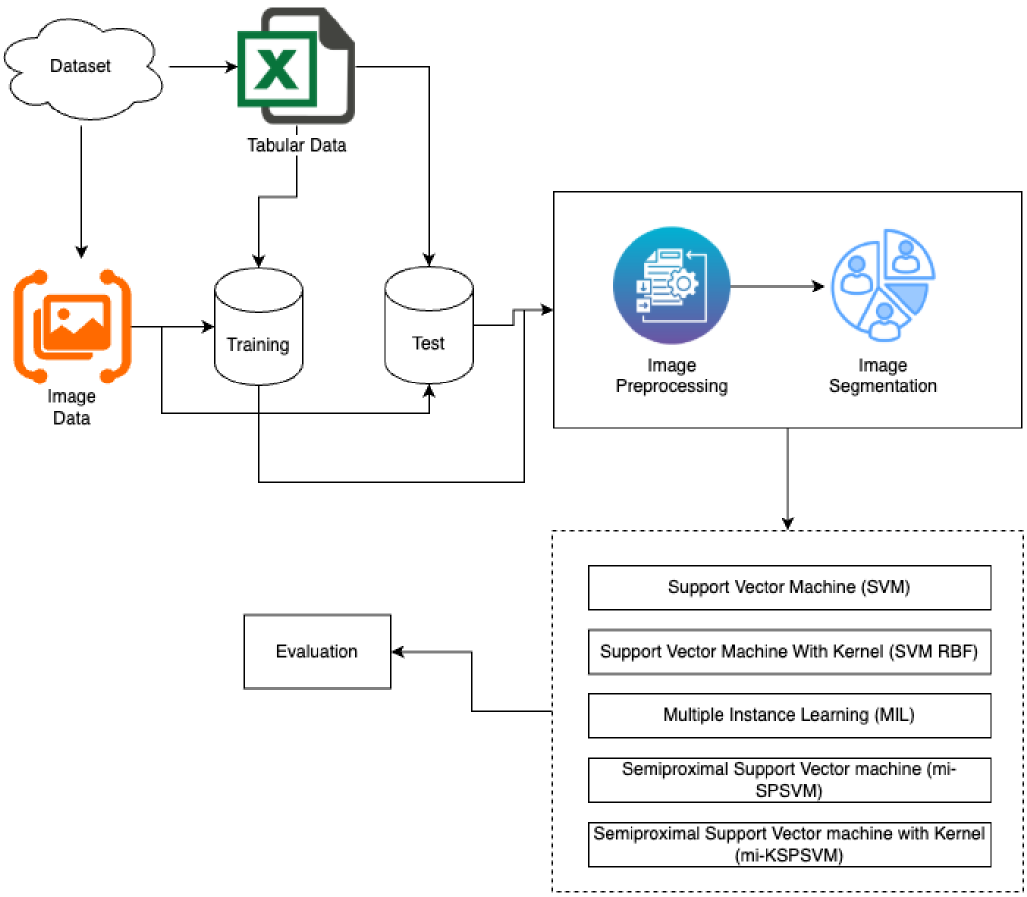

Figure 2. Overall Pipeline.

3.1. Data Acquisition

This study utilized the MESSIDOR dataset [22], a well-established resource for retinal image analysis. It contains 1,200 color fundus images collected from three French eye clinics. The images are labeled with expert-reviewed annotations indicating DR severity and macular edema risk. Among them, 898 images have a resolution of 1440×960 pixels, while 302 are 2240×1488, all in 8-bit RGB format. These properties make MESSIDOR suitable for segmentation and classification tasks in medical imaging.

3.2. Data Preprocessing

Preprocessing is the most common step for each segment, involving noise removal and contrast enhancement. In addition to improving contrast and reducing noise, preprocessing may also include normalization and image resizing to improve overall processing accuracy.

The preprocessing process starts with image resizing. During the segmentation process, high-resolution images consume a significant amount of time. So, image resizing is required. To resize the images, we used a popular "Matlab" function called "imresize" with a scaling of 25% and the "lanczos3" method.

After resizing the images, we extract the green image from the RGB channel. However, the choice of the green channel image is not a rule. But from the anatomical structure of interest, green channel images are preferred for segmentation. This green channel has a balance between the red and blue channels, which helps reduce noise and enhance the contrast between different structures in the image.

After resizing, the next step is to create a binary mask, which removes the ineffective area taken by the eye fundus camera. To obtain a binary mask, Otsu's thresholding method is applied with a threshold value of $T \leq 0.05$. Subsequently, Contrast Limited Adaptive Histogram Equalization (CLAHE) is employed to improve the vessel visibility in the image [13]. CLAHE enhances image contrast by applying a "ClipLimit" to small regions known as tiles instead of the entire image. After processing, the neighboring tiles are seamlessly merged using bilinear interpolation. This method is explicitly applied to the green channel of retinal images using carefully selected parameter values. To apply CLAHE in this study, we used a Matlab function named "adapthisteq" with (8, 8) tiles, default (0.01) cliplimit, and uniform distribution.

3.3. Algorithm

This research is divided into two parts: one involves image segmentation using traditional image processing techniques, and the other utilizes optimized algorithms to classify images based on their labels.

3.3.1. Multiple Instance Learning

Multiple Instance Learning (MIL) is a technique where data is structured into groups, referred to as "bags," each containing several elements called "instances". In this setting, only the labels for the bags are provided during training. The primary goal of MIL is to classify these bags based on the collective properties of their instances. In the linear separation, points (positive and negative) are separated by a simple linear line (shown in the Figure 3).

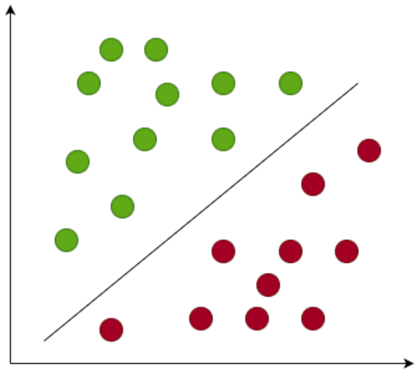

Figure 3. *Linear Separation.*

However, because of the complexity of the data, it often failed to generalize. To achieve generalization capability, the SVM model emphasizes two key factors: maximizing the margin (enhancing generalization) and minimizing the classification error (as shown in Figure 4).

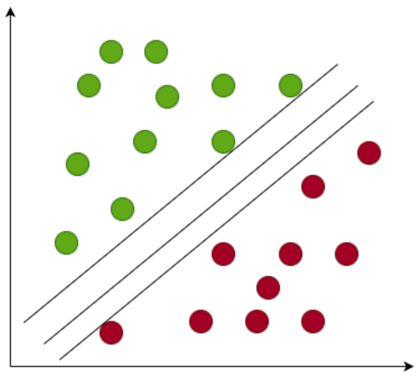

Figure 4. *SVM Hyperplane Separation.*

This study conducted four optimized MIL algorithms for the image classification task, which are: a) MIL-RL, b) SVM and SVM (RBF), c) mi-SPSVM, and d) mi-KSPSVM.

3.3.2. MIL-RL

Multiple Instance Learning with Lagrangian relaxation, in short MIL-RL, is an approach that combines MIL and Lagrangian relaxation [10]. MIL-RL is the following (1):

$$\begin{cases} \min\limits_{w,b,y,\epsilon,\psi} & \frac{1}{2}||w||^2 + C\sum_{i=1}^{m}\sum_{j\in J_i^+}\epsilon_j + C\sum_{i=1}^{k}\sum_{j\in J_i^-}\psi_j \\ & \epsilon_j \geq 1 - y_j(w^T x_j + b) \quad j \in J_i^+, i = 1, ..., m \\ & \psi_j \geq 1 + (w^T x_j + b) \quad j \in J_l^-, l = 1, ..., k \\ & \sum_{j\in J_i^+}\frac{y_j+1}{2} \geq 1, i = 1...m \\ & y_j \in \{-1, +1\}, \quad j \in J_i^+, i = 1...m \\ & \epsilon_j \geq 0 \quad j \in J_i^+, i = 1, ..., m \\ & \psi_j \geq 0 \quad j \in J_l^-, l = 1, ..., k \end{cases} \tag{1}$$

3.3.3. mi-SPSVM

Another classification technique that utilizes the standard MIL assumption to solve binary cases is called the Semiproximal SVM (SP-SVM). Specifically, this algorithm is a representation of a well-known MIL SVM, a heuristic algorithm called mi-SVM has been proposed [9]. mi-SPSVM is the following (2):

$$\begin{cases} \min\limits_{w,b,\epsilon}\frac{1}{2}(w^T w + b^2) + \frac{C}{2}\sum_{i=1}^{m}\sum_{j\in J_i^+}||\epsilon_j||^2 + C\sum_{i=1}^{k}\sum_{j\in J_i^-}\psi_j \\ \epsilon_j = 1 - (w^T x_j + b) \quad j \in J_i^+, i = 1, ..., m \\ \psi_j \geq 1 + (w^T x_j + b) \quad j \in J_l^-, l = 1, ..., k \\ \psi_j \geq 0 \quad j \in J_l^-, l = 1, ..., k \end{cases} \tag{2}$$

Three terms make up the objective function of problem (2): the first one maximises the margin concerning w and b. The Second one maximises the “proximity” of the positive bag instances. The final one minimises the misclassification error of the negative bag instances.

3.3.4. mi-KSPSVM

To work with both efficiency and generalization capability, [8] presents a kernel-based semiproximal SVM for MIL (mi-KSPSVM). Optimization algorithms are primarily associated with three types of kernels: linear, Gaussian (RBF), and Hyperbolic Tangent. Among those, this optimization algorithm (mi-KSPSVM) is connected with the RFB kernel to capture and calculate the nonlinear structure from the data. mi-KSPSVM is the following (3):

$$\begin{cases} \min\limits_{\lambda,\mu} & \frac{1}{2}\Bigg(\sum_{i=1}^{m}\sum_{j\in J_i^+}\lambda_i\lambda_j\phi(x_i)^T\phi(x_j) \\ & - 2\sum_{i=1}^{m}\sum_{l=1}^{k}\lambda_i\mu_l\phi(x_i)^T\phi(x_l) \\ & + \sum_{l=1}^{k}\sum_{j\in J_l^-}\mu_l\mu_j\phi(x_l)^T\phi(x_j)\Bigg) \\ & + \frac{1}{2}\Bigg(\sum_{i=1}^{m}\sum_{j\in J_i^+}\lambda_j - \sum_{l=1}^{k}\sum_{j\in J_i^-}\mu_j\Bigg)^2 \\ & - \frac{\Big(\sum_{i=1}^{m}\sum_{j\in J_i^+}\lambda_j\Big)^2}{2C} + \sum_{i=1}^{m}\sum_{j\in J_i^+}\lambda_j + \sum_{l=1}^{k}\sum_{j\in J_i^-}\mu_j \\ & 0 \leq \mu_j \leq C, \quad j \in J_i^-, i = 1, ..., k \end{cases} \tag{3}$$

3.3.5. Image Segmentation

The retinal vascular network provides valuable insights into the eye's health condition. It stands as the sole non-invasive technique capable of visualizing blood vessels within the human body. In retinal imaging, blood vessels are among the most critical anatomical structures for identifying early signs of diabetic retinopathy (DR). Accurate segmentation of these vessels is crucial for supporting the diagnosis of systemic vascular disorders. Retinal vessels appear as a combination of fine and coarse lines, with thicker vessels being more easily detectable than the thinner ones. This work addresses the segmentation of both fine and thick vessels. For isolating thin vessels from fundus images, the Frangi filter [18] has proven to be a reliable approach. Meanwhile, morphological operations are also effective for identifying and extracting the thicker vessel structures.The pseudocode 1 of vascular segmentation is given below:

Algorithm 1 Blood Vessel Segmentation

Require: $Image(x, y), Mask(x, y)$
Ensure: $Output(x, y)$
Thick Vessel Extraction
$R(x, y), Mean \leftarrow RemoveBlackBackground(Image, Mask)$
$I_{thick}(x, y) \leftarrow ThickVessel(R, Mean, Mask)$ ▷ Output of Thick Blood Vessel
Thin Vessel Extraction
$I_2(x, y) \leftarrow FrangiFilter(MedFilter(Image))$ ▷ Output of Thin Blood Vessel
$I_{thin}(x, y) \leftarrow AdaptiveOtsu(I_2)$
Final Image
$O(x, y) \leftarrow max(I_{thick}, I_{thin})$

3.4. Evaluation

This section outlines the evaluation process for models. To evaluate the model's performance in an unseen scenario, the following matrices were calculated:

1. *Testing Correctness (TC):* The assessment of the model's accuracy in accurately identifying medical images is called "testing correctness in medical image classification." It involves comparing the test dataset's ground-truth labels with the predicted labels generated by the classification model. The equation of testing correctness is:

$$\text{TC} = \frac{\text{\# of correctly classified in the testing set}}{\text{\# total points in the testing set}}$$

2. *Sensitivity:* Sensitivity represents the ratio of positive images correctly classified to the actual positive images. It evaluates how well the model detects positive examples and is often referred to as sensitivity. To calculate the sensitivity of a model, the following:

$$\text{Sensitivity} = \frac{\text{\# of positive images correctly classified}}{\text{\# actual positive images}}$$

3. *Specificity:* Specificity calculates the proportion of correctly predicted negative (true negative) images among all the actual negative images. It assesses the model's ability to identify negative cases accurately. To calculate the specificity of a model, the following:

$$\text{Specificity} = \frac{\text{\# of negative images correctly classified}}{\text{\# actual negative images}}$$

4. *Precision:* Precision measures the fraction of accurate optimistic predictions out of all instances that the model has labeled as positive. Reflects the model's accuracy in identifying positive cases.

$$\text{Precision} = \frac{\text{\# of positive images correctly classified}}{\text{\# total points classified as positive}}$$

5. *F1-Score:* F1 Score is the inverse of the arithmetic mean of precision and sensitivity. Provides a balanced measure that considers both precision and recall.

$$\text{F1-Score} = 2 \times \frac{\text{sensitivity} \times \text{precision}}{\text{sensitivity} + \text{precision}}$$

4. Result and Discussion

The experimental results in this study are conducted based on two approaches. One is without preprocessing on the images, while the other is with preprocessed images. To verify any significant difference between the results of these two tests, we run our models using a prepared dataset with 5-fold, 10-fold, and LOOCV cross-validation, and obtain numerical results. After presenting the results from the above approaches, we primarily considered sensitivity because it provides information on whether an image is correctly classified. We also evaluated testing correctness, specificity, F1 scores, and CPU Time. All the numerical results are described below.

4.1. Without Preprocessed Image

The first experiment on diabetic retinopathy disease was conducted without image processing, where a full image is considered and applied to all five models, which is later used for inference. Table 1 illustrates experiment results.

Table 1. *Performance metrics of the fundus image without preprocessing.*

Cross Validation	Model	Testing Correctness	Sensitivity	Specificity	F1-Score	CPU Time(s)
	mi-KSPSVM	55.50	78.32	33.13	63.52	0.27
	mi-SPSVM	53.50	74.47	33.41	61.40	0.67
5-Fold	MIL-RL	49.50	8.21	91.06	11.64	9.33
	SVM	47.50	35.54	57.76	38.88	1.51
	SVM (RBF)	58.50	62.64	54.17	59.93	0.28
	mi-KSPSVM	55.50	76.18	33.24	62.69	0.18
	mi-SPSVM	54.00	76.07	32.60	61.61	0.33
10-Fold	MIL-RL	48.00	12.66	86.67	10.38	14.94
	SVM	50.00	43.58	56.63	45.06	1.57
	SVM (RBF)	57.00	62.15	54.04	58.49	0.09
	mi-KSPSVM	52.00	79.00	25.00	62.20	0.36
	mi-SPSVM	54.50	74.00	35.00	61.92	2.84
LOO	MIL-RL	39.50	52.00	27.00	46.22	20.94
	SVM	47.50	38.00	57.00	41.99	1.64
	SVM (RBF)	55.50	58.00	53.00	56.59	0.02

4.2. With Preprocessed Image

After evaluating the model using images without preprocessing, the second step was to repeat the process with preprocessed images. This step exactly replicated the previous work. Table 2 illustrate experiment results.

Table 2. *Performance metrics of the fundus image with preprocessing.*

Cross Validation	Model	Testing Correctness	Sensitivity	Specificity	F1-Score	CPU Time
	mi-KSPSVM	57.50	71.57	44.03	62.17	0.23
	mi-SPSVM	66.00	68.37	63.82	66.69	0.27
5-Fold	MIL-RL	59.50	40.64	79.13	43.74	3.61
	SVM	59.50	55.99	65.53	56.87	1.51
	SVM (RBF)	58.00	58.00	51.88	60.40	0.21
	mi-KSPSVM	59.00	72.80	44.61	62.92	0.47
	mi-SPSVM	63.50	66.87	60.70	64.30	0.42
10-Fold	MIL-RL	68.00	58.10	77.78	59.39	4.82
	SVM	56.50	48.12	68.06	50.47	1.44
	SVM (RBF)	55.50	61.21	48.83	57.01	0.10
	mi-KSPSVM	58.00	84.00	32.00	66.67	0.17
	mi-SPSVM	68.00	69.00	67.00	68.32	1.56
LOO	MIL-RL	55.50	33.00	78.00	42.58	7.17
	SVM	51.00	36.00	66.00	42.35	1.52
	SVM (RBF)	59.00	68.00	50.00	62.39	0.01

It was observed that image preprocessing led to a consistent improvement in classification accuracy and F1-score across most models. For instance, the mi-SPSVM model under 5-Fold cross-validation improved its testing correctness from 53.50% to 66.00% and F1-Score from 61.40% to 66.69% after preprocessing. Similarly, the MIL-RL model showed a substantial increase in testing correctness from 48.00% to 68.00% under 10-Fold validation, though at the cost of longer computation time.

Preprocessing also influenced the balance between sensitivity and specificity. While models like mi-KSPSVM (LOO) demonstrated a high sensitivity of 84.00%, the specificity dropped to 32.00%. On the other hand, MIL-RL consistently exhibited high specificity (up to 91.06% without preprocessing), but its sensitivity remained low, making it overly conservative for practical applications. Regarding computational efficiency, the SVM (RBF) model showed both stability in performance and low CPU time, making it favorable for real-time applications. Notably, the SVM (RBF) model under LOO validation improved its F1-score from 56.59% to 62.39% after preprocessing, while reducing CPU time from 0.02s to 0.01s.

Overall, the results demonstrate that preprocessing significantly enhances the performance of classification models in fundus image analysis. With a testing correctness of 68.00% and an F1-score of 68.32%, mi-SPSVM with preprocessing under LOO cross-validation performed the best overall out of all configurations. This suggests that appropriate preprocessing not only enhances model accuracy but also contributes to more balanced and reliable diagnostic output in medical image segmentation tasks.

4.3. Comparative Analysis

Additionally, a qualitative analysis of the classifications related to different optimization models has been conducted. The bar graphs of Figures 5, 6, 7 present a comparative analysis of the generated graphs, showing that preprocessing consistently improves sensitivity across all models and validation schemes.

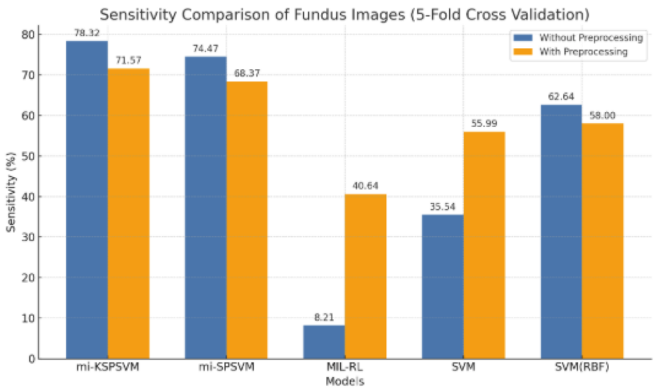

Figure 5. *5-fold CV Results.*

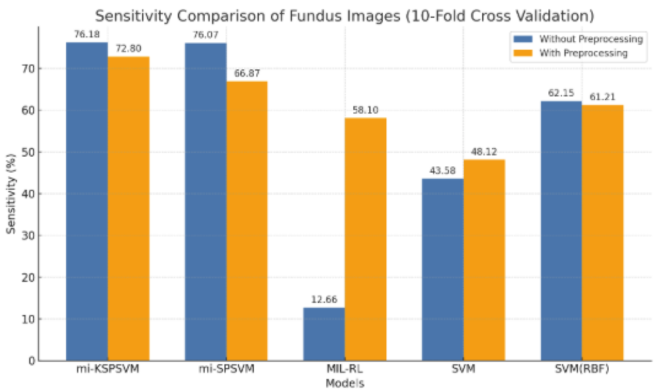

Figure 6. *10-fold CV Results.*

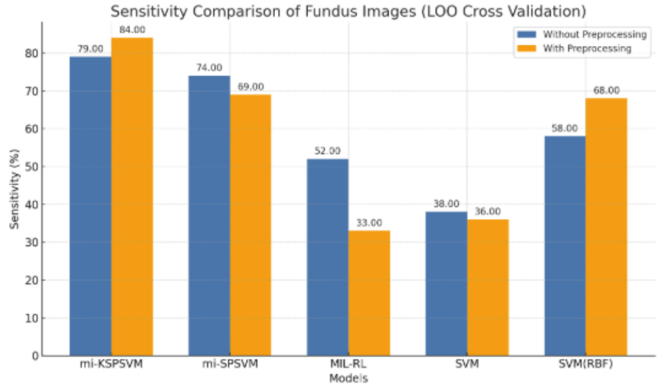

Figure 7. *LOO-CV Results.*

mi-KSPSVM achieves the highest sensitivity under LOO-CV (84%), while MIL-RL shows the most significant improvement under 10-Fold CV. SVM (RBF) demonstrates stable performance with minimal fluctuation and the lowest computational cost. In contrast, MIL-RL lags in sensitivity despite preprocessing. Overall, mi-SPSVM and mi-KSPSVM emerge as the most sensitive and reliable models post-preprocessing.

4.4. Discussion

This study focused on image classification using segmented retinal fundus images, where the segmentation process was carried out using traditional image processing techniques. This approach offered computational efficiency while maintaining acceptable anatomical integrity of key retinal structures, such as vessels. The method remained lightweight and practical for implementation on systems with limited computational resources.

All experiments were performed on a Mac OS machine equipped with an M1 processor and 8GB of RAM, demonstrating that it is possible to obtain consistent and meaningful results without high-end hardware.

Although image processing techniques may lack flexibility in handling diverse pathological variations, they offer faster execution, lower memory consumption, and more straightforward implementation compared to more complex modeling techniques. The results confirm that a segmentation-first pipeline can effectively support retinal image classification tasks, especially in environments with limited hardware availability.

5. Conclusion

This work focused on identifying healthy and diseased cases from retinal fundus images through a two-step approach. The primary objective was to enhance classification performance by extracting the relevant features from the image, rather than using the entire image directly. The classification was conducted using an MIL framework, which proved effective in analyzing medical images by learning global labels from localized instances. This suggests that further refinement in the segmentation process could yield better results. Future work will focus on refining segmentation techniques to eliminate irrelevant regions and enhance classification performance, utilizing the latest deep learning approaches, particularly in the segmentation tasks.

ORCID

0009-0000-7906-1823 (Mahbub Hasan)

Abbreviations

DRD	Diabetic Retinopathy Disease
DME	Diabetic Macular Edema
OD	Optic Disk
BV	Blood Vessel
DR	Diabetic Retinopathy
TC	Testing Correctness
CV	Cross Validation
LOO	Leave-One-Out
MIL	Multiple Instance Learning
SVM	Support Vector Machine
MIL-RL	Multiple Instance Learning With Lagrangian Relaxation
mi-SPSVM	Multiple Instance, Semiproximal Support Vector Machine
CLAHE	Contrast Limited Adaptive Histogram Equalization

Author Contributions

Mahbub Hasan is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Ahmad Fadzil, M. H., Izhar, L. I., Venkatachalam, P. A., & Karunakar, T. V. N. (2007). Extraction and reconstruction of retinal vasculature. Journal of Medical Engineering & Technology, 31(6), 435-442. https://doi.org/10.1080/03091900601111201

[2] Avolio, M., Fuduli, A., Vocaturo, E. and Zumpano, E., 2023, May. On detection of diabetic retinopathy via multiple instance learning. In Proceedings of the 27th International Database Engineered Applications Symposium (pp. 170-176). https://doi.org/10.1145/3589462.3589490

[3] M. Mateen, J. Wen, M. Hassan, N. Nasrullah, S. Sun and S. Hayat, "Automatic Detection of Diabetic Retinopathy: A Review on Datasets, Methods and Evaluation Metrics," in IEEE Access, vol. 8, pp. 48784-48811, 2020, https://doi.org/10.1109/ACCESS.2020.2980055

[4] Li, T., Bo, W., Hu, C., Kang, H., Liu, H., Wang, K., & Fu, H. (2021). Applications of deep learning in fundus images: A review. Medical image analysis, 69, 101971. https://doi.org/10.1016/j.media.2021.101971

[5] Andrews, S., Tsochantaridis, I. and Hofmann, T., 2002. Support vector machines for multiple-instance learning. Advances in neural information processing systems, 15. https://doi.org/10.5555/2968618.2968690

[6] A. Astorino, A. Fuduli, P. Veltri, and E. Vocaturo, "On a recent algorithm for multiple instance learning. Preliminary applications in image classification," 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 2017, pp. 1615-1619, https://doi.org/10.1109/BIBM.2017.8217901

[7] Astorino, A., Fuduli, A., Gaudioso, M. and Vocaturo, E., 2018, June. A multiple instance learning algorithm for color images classification. In Proceedings of the 22nd International Database Engineering & Applications Symposium (pp. 262-266). https://doi.org/10.1145/3216122.3216144

[8] Avolio, M. and Fuduli, A., 2024. The semiproximal SVM approach for multiple instance learning: a kernel-based computational study. Optimization Letters, 18(2), pp. 635-649. https://doi.org/10.1007/s11590-023-02022-8

[9] Avolio, M. and Fuduli, A., 2020. A semiproximal support vector machine approach for binary multiple instance learning. IEEE transactions on neural networks and learning systems, 32(8), pp. 3566-3577. https://doi.org/10.1109/TNNLS.2020.3015442

[10] Astorino, A., Fuduli, A. and Gaudioso, M., 2019. A Lagrangian relaxation approach for binary multiple instance classification. IEEE transactions on neural networks and learning systems, 30(9), pp. 2662-2671. https://doi.org/10.1109/TNNLS.2018.2885852

[11] Y. Xiao, B. Liu and Z. Hao, "Multi-Instance Nonparallel Tube Learning," in IEEE Transactions on Neural Networks and Learning Systems, vol. 36, no. 2, pp. 2563-2577, Feb. 2025, https://doi.org/10.1109/TNNLS.2023.3347449

[12] Zumpano, E., Fuduli, A., Vocaturo, E. and Avolio, M., 2021, July. Viral pneumonia images classification by Multiple Instance Learning: preliminary results. In Proceedings of the 25th International Database Engineering & Applications Symposium (pp. 292-296). https://doi.org/10.1145/3472163.3472170

[13] Jeong Y, Hong Y-J, Han J-H. Review of Machine Learning Applications Using Retinal Fundus Images. Diagnostics. 2022; 12(1): 134. https://doi.org/10.3390/diagnostics12010134

[14] Khan, K. B., Khaliq, A. A., Jalil, A. et al. A review of retinal blood vessels extraction techniques: challenges, taxonomy, and future trends. Pattern Anal Applic 22, 767-802 (2019). https://doi.org/10.1007/s10044-018-0754-8

[15] Sarki, R., Ahmed, K., Wang, H. et al. Image Preprocessing in Classification and Identification of Diabetic Eye Diseases. Data Sci. Eng. 6, 455-471 (2021). https://doi.org/10.1007/s41019-021-00167-z

[16] Edupuganti, V. G., Chawla, A., & Kale, A. (2018). Automatic Optic Disk and Cup Segmentation of Fundus Images Using Deep Learning. 2227-2231. https://doi.org/10.1109/icip.2018.8451753

[17] A. Khawaja, T. M. Khan, K. Naveed, S. S. Naqvi, N. U. Rehman and S. Junaid Nawaz, "An Improved Retinal Vessel Segmentation Framework Using Frangi Filter Coupled With the Probabilistic Patch Based Denoiser," in IEEE Access, vol. 7, pp. 164344-164361, 2019, https://doi.org/10.1109/ACCESS.2019.2953259

[18] Frangi, A. F., Niessen, W. J., Vincken, K. L., Viergever, M. A. (1998). Multiscale vessel enhancement filtering. In: Wells, W. M., Colchester, A., Delp, S. (eds) Medical Image Computing and Computer-Assisted Intervention - MICCAI'98. MICCAI 1998. Lecture Notes in Computer Science, vol 1496. Springer, Berlin, Heidelberg. https://doi.org/10.1007/BFb0056195

[19] Mahapatra, S., Agrawal, S., Mishro, P. K., & Pachori, R. B. (2022). A novel framework for retinal vessel segmentation using optimal improved frangi filter and adaptive weighted spatial FCM. Computers in biology and medicine, 147, 105770. https://doi.org/10.1016/j.compbiomed.2022.105770

[20] Lara Rodríguez, Luis David, & Urcid Serrano, Gonzalo. (2016). Exudates and Blood Vessel Segmentation in Eye Fundus Images Using the Fourier and Cosine Discrete Transforms. Computación y Sistemas, 20(4), 697-708. https://doi.org/10.13053/cys-20-4-2305

[21] Kumar, Ashish & Huang, Yo-Ping. (2021). Detection of Retinal Blood Vessels with their Geometrical Physical Characteristics for Retinopathy of Prematurity (ROP).

[22] Decencière, Etienne and Zhang, Xiwei and Cazuguel, Guy and Lay, Bruno and Cochener, Béatrice and Trone, Caroline and Gain, Philippe and Ordóñez-Varela, John-Richard and Massin, Pascale and Erginay, Ali, 2014. Feedback on a publicly distributed image database: the Messidor database. Image Analysis & Stereology, pp.231-234. https://doi.org/10.5566/ias.1155

[23] Odeh, M. Alkasassbeh and M. Alauthman, "Diabetic Retinopathy Detection using Ensemble Machine Learning," 2021 International Conference on Information Technology (ICIT), 2021, pp. 173-178, https://doi.org/10.1109/ICIT52682.2021.9491645

[24] Carbonneau, M. A., Cheplygina, V., Granger, E. and Gagnon, G., 2018. Multiple instance learning: A survey of problem characteristics and applications. Pattern recognition, 77, pp. 329-353. https://doi.org/10.1016/j.patcog.2017.10.009