informatics

MDPI

Article

An Optimized Superpixel-Based Framework for Hard Exudates Segmentation Incorporating Multi-Stage Preprocessing

Kittipol Wisaeng [1,*] and Sonthinee Waiyarat [2]

[1] Mahasarakham University Research Laboratory of Blockchain and Artificial Intelligence for Interdisciplinary Innovation, Mahasarakham Business School, Mahasarakham University, Mahasarakham 44150, Thailand
[2] College of Creative Agriculture for Society, Srinakharinwirot University, Bangkok 26120, Thailand; sonthinee@g.swu.ac.th
* Correspondence: kittipol.w@acc.msu.ac.th

Abstract

Hard exudates (HEs) are among the earliest clinical manifestations of diabetic retinopathy (DR) and serve as important indicators for disease diagnosis and progression assessment. Accurate segmentation of HEs is therefore essential for early intervention and prevention of vision loss. However, automatic HEs detection remains challenging because lesion appearance, illumination conditions, and anatomically similar retinal structures such as the optic disc vary. This study proposes a knowledge-guided hard exudate detection framework based on a Superpixel-Based Framework (SBF) combined with adaptive feature analysis. The proposed method integrates retinal image enhancement, CIELAB color-space transformation, SBF superpixel generation, adaptive luminance–chromaticity thresholding, optic disc suppression, and morphological refinement to improve lesion localization and reduce false-positive detections. Unlike deep learning approaches that require large annotated datasets and substantial computational resources, the proposed framework leverages perceptually meaningful color information and retinal anatomical knowledge to achieve robust, interpretable segmentation. We evaluated the framework on the publicly available DIARETDB1 and STARE datasets and compared it with K-means Clustering, Fuzzy C-Means (FCM), Mean Shift Segmentation, Region Growing, U-Net, and Convolutional Neural Network (CNN) based segmentation methods. The proposed framework achieved the best performance on the Standard Diabetic Retinopathy Database (DIARETDB1), with 98.24% accuracy, 95.81% sensitivity, 99.12% specificity, a Dice coefficient of 95.58%, and an IoU of 91.42%. Similarly, on the STructured Analysis of the Retina (STARE) dataset, the method achieved 97.89% accuracy and 90.21% IoU. The results demonstrate that the proposed method provides accurate, computationally efficient, and interpretable HEs segmentation for automated DR screening systems.

Keywords: diabetic retinopathy; hard exudates; retinal fundus image; superpixel-based framework segmentation; CIELAB color space; optic disc suppression; image segmentation; computer-aided diagnosis

Check for updates

Academic Editor: Luc Bidaut

Received: 19 June 2026
Revised: 9 September 2026
Accepted: 13 September 2026
Published: 17 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Diabetic retinopathy (DR) is one of the leading causes of preventable blindness worldwide and poses a significant public health challenge, particularly in developing countries where access to regular eye screening is limited. Among the early clinical manifestations of DR, HEs are lipid and protein deposits that appear as bright yellow lesions in retinal

fundus images. These lesions are critical biomarkers for diagnosing and monitoring DR progression, as their presence and distribution are strongly associated with retinal edema and vision impairment. Therefore, accurate and automated detection of HEs is essential for early diagnosis and timely treatment, ultimately reducing the risk of vision loss [1]. Traditional HEs detection methods rely heavily on manual inspection by ophthalmologists, which is time-consuming, subjective, and prone to inter-observer variability. To address these limitations, researchers have extensively developed computer-aided diagnosis (CAD) systems to help clinicians analyze retinal images. Early approaches primarily employed classical image processing techniques such as thresholding, morphological operations, and clustering algorithms, including k-means and FCM. Although these methods demonstrated promising results, they often suffer from sensitivity to illumination variations, low contrast, and noise inherent in fundus images, leading to suboptimal segmentation performance [2]. In recent years, machine learning and deep learning approaches have significantly advanced retinal image analysis. Convolutional neural networks (CNN) and hybrid models have achieved high accuracy in detecting DR lesions, including HEs. However, these approaches typically require large annotated datasets, high computational resources, and extensive training time, which may limit their applicability in real-world clinical settings, especially in resource-constrained environments [3]. Furthermore, deep learning models often lack interpretability, making it difficult for clinicians to fully trust their predictions in critical diagnostic scenarios. To overcome these challenges, superpixel-based segmentation has emerged as an effective alternative for medical image analysis. Superpixel groups pixels into perceptually meaningful regions, preserving boundary information while reducing computational complexity. Among various superpixel algorithms, SBF has gained popularity for its efficiency, simplicity, and ability to generate compact, uniform regions. By operating on superpixels rather than individual pixels, segmentation algorithms can achieve improved robustness and accuracy in identifying pathological regions such as HEs. Despite the advantages of superpixel-based methods, segmentation quality largely depends on effective preprocessing. Retinal fundus images often exhibit uneven illumination, low contrast, and noise artifacts, which can significantly degrade segmentation performance. Therefore, a robust multi-stage preprocessing pipeline is essential to enhance image quality before segmentation. Techniques such as multi-level normalization and contrast enhancement have been widely adopted to improve the visibility of retinal structures and lesions [4]. In this study, we propose an optimized framework for HEs segmentation in retinal fundus images that combines superpixel-based methods with multi-stage image preprocessing. The proposed approach integrates five normalization techniques, contrast enhancement, and noise removal to effectively address common challenges in retinal imaging.

The remainder of this paper is structured as follows: Section 2 presents a comprehensive review of related studies on HEs segmentation, diabetic retinopathy screening, retinal image preprocessing, and superpixel-based image analysis. Section 3 details the proposed knowledge-guided SBF for HEs detection. Section 4 describes the experimental design, benchmark datasets, evaluation metrics, and implementation settings. Section 5 presents the quantitative and qualitative results, comparative analyses, and ablation studies. Finally, Section 6 discusses the findings, practical implications, and limitations of the proposed approach.

2. Literature Review

Automated detection of HEs in retinal fundus images has been extensively studied because of its clinical importance in the early diagnosis of DR. Traditional image processing approaches have laid the foundation for HEs segmentation by using intensity-based thresholding, region growing, and morphological operations. Early studies applied global

and adaptive thresholding techniques to identify bright lesions; however, these methods were highly sensitive to illumination variation and often misclassified optic disc regions as exudates [5]. To improve segmentation performance, researchers introduced clustering-based techniques such as k-means and FCM, enabling pixel grouping based on intensity and color features. Although FCM demonstrated improved flexibility in handling uncertainty, it remained sensitive to noise and initialization parameters [6]. Subsequently, researchers explored feature-based and machine learning approaches for HEs detection. Methods incorporating handcrafted features such as color, texture, and edge descriptors, combined with classifiers such as support vector machines (SVMs) and random forests, showed improved detection accuracy [7]. These approaches achieved better generalization than purely threshold-based methods but required careful feature engineering and domain expertise. Moreover, their performance often degraded on heterogeneous datasets with varying image quality and acquisition conditions. In recent years, deep learning techniques, particularly CNNs, have significantly advanced retinal image analysis. CNN-based models have achieved state-of-the-art performance in detecting DR lesions, including HEs, by automatically learning hierarchical feature representations from large-scale datasets [8]. Variants such as U-Net and fully convolutional networks (FCNs) have been widely used for pixel-wise segmentation tasks. Despite their success, these methods require extensive annotated data, high computational resources, and long training times. Additionally, the "black-box" nature of deep learning models limits interpretability, a critical concern in medical applications.

To address these limitations, superpixel-based segmentation has emerged as a promising alternative. Superpixel algorithms group pixels into perceptually meaningful regions, preserving structural boundaries while reducing computational complexity. The SBF algorithm is particularly popular due to its efficiency and ability to produce compact, uniform superpixels. Studies have shown that superpixel-based methods can improve lesion localization and reduce noise sensitivity compared with pixel-wise approaches [9]. However, their effectiveness depends strongly on input image quality and the preprocessing steps applied. Image preprocessing plays a crucial role in improving segmentation performance in retinal analysis. Techniques such as contrast-limited adaptive histogram equalization (CLAHE), illumination correction, and noise filtering have been widely used to enhance the visibility of retinal structures. Multi-stage preprocessing pipelines that combine normalization, contrast enhancement, and denoising have significantly improved detection of small, low-contrast lesions, such as HEs [10]. Nevertheless, existing studies often use limited preprocessing strategies and do not fully explore integrating multiple normalization techniques within a unified framework. Despite the significant progress in HEs detection, several limitations remain evident in the literature. First, traditional and clustering-based methods are sensitive to noise, illumination variations, and low contrast. Second, while deep learning approaches achieve high accuracy, they require large annotated datasets, incur high computational costs, and lack interpretability, making them less suitable for real-time or resource-constrained environments. Third, although superpixel-based segmentation balances efficiency and accuracy, existing methods often rely on limited or suboptimal preprocessing pipelines. In particular, comprehensive frameworks that integrate multilevel normalization, contrast enhancement, and noise removal to fully optimize input image quality before segmentation are lacking. Furthermore, few studies systematically combine these preprocessing techniques with superpixel-based segmentation to improve robustness across diverse datasets. Table 1 summarizes and compares representative HEs detection and segmentation methods reported in the literature.

Table 1. Comparative analysis of existing HEs detection and segmentation methods.

Ref.	Method	Dataset	Main Technique	Strengths	Limitations	Research Gap
[1]	Hard Exudate Detection and Maculopathy Grading	Fundus Images	Multi-stage HEs Detection	Integrates HEs detection and disease grading	Focuses primarily on grading performance rather than precise lesion segmentation	Limited emphasis on superpixel-guided lesion localization
[2]	FCM Clustering	Retinal Images	Fuzzy C-Means	Handles uncertain boundaries effectively	Sensitive to initialization and noise	Limited robustness in low-contrast retinal images
[6]	K-means Clustering	Fundus Images	Color-based Clustering	Computationally efficient	Hard clustering causes boundary inaccuracies	Poor adaptation to heterogeneous lesion appearance
[4]	U-Net	Fundus Images	Deep CNN Segmentation	High segmentation accuracy	Requires large annotated datasets	High computational requirements and limited interpretability
[7]	SVM + Deep Learning	Fundus Images	Object Detection Framework	Accurate lesion localization	Requires extensive training data	Detection-oriented rather than pixel-level segmentation
[5]	Intensity Trend Analysis	Fundus Images	Incremental-Decremental Intensity Modeling	Effective lesion localization	Sensitive to illumination variation	Limited integration with anatomical constraints
Proposed Method	DIARETDB1, STARE	CLAHE + SLIC + Adaptive Thresholding + Optic Disc Suppression + Morphological Refinement	An interpretable multi-stage framework combining color enhancement, superpixel analysis, lesion candidate selection, and anatomical suppression	High Accuracy (98.24%), Dice (95.58%), and IoU (91.42%); computationally efficient; interpretable; requires no large-scale training dataset	Performance may depend on parameter selection and image quality	Addresses limitations of clustering and deep learning methods by integrating superpixel-guided lesion segmentation with optic disc suppression and adaptive refinement

A comparative analysis of existing HEs detection and segmentation methods reported in the literature is presented in Table 1. Traditional clustering-based approaches, such as K-means and FCM, offer computational simplicity but often suffer from noise sensitivity and inaccurate lesion boundary delineation. Deep learning-based methods, including U-Net and CNN architectures, achieve higher segmentation accuracy but generally require large annotated datasets and substantial computational resources. In contrast, the proposed framework integrates image enhancement, SBF segmentation, adaptive lesion candidate selection, optic disc suppression, and morphological refinement into a unified, interpretable pipeline. This combination enables accurate localization of HEs while maintaining computational efficiency and robustness across multiple retinal image datasets. The proposed method consists of three key stages:

1. Multi-Stage Preprocessing: The input retinal fundus images undergo a comprehensive preprocessing pipeline that includes five normalization techniques to standardize intensity variations, followed by contrast enhancement using CLAHE and noise reduction using hybrid filtering (median and Gaussian filters).
2. Superpixel-Based Segmentation: The preprocessed images are segmented using the SBF algorithm, which groups pixels into meaningful regions while preserving boundary information. Feature extraction is performed at the superpixel level, incorporating color and intensity characteristics to identify candidate HEs regions.
3. Post-Processing and Refinement: Adaptive thresholding and morphological operations are applied to refine segmentation results, remove false positives (e.g., optic disc regions), and generate the final HEs mask.

3. Research Methodology

The proposed framework is evaluated on publicly available retinal datasets, including DIARETDB1 and STARE, containing annotated retinal images with HEs. A multi-stage preprocessing pipeline applies five normalization techniques, contrast enhancement using CLAHE, and noise removal via median and Gaussian filtering. Superpixel-based segmentation uses SBF to group pixels into homogeneous regions and extract candidate lesions. Finally, post-processing involves adaptive thresholding and morphological operations to refine segmentation, eliminate false positives (e.g., optic disc), and generate accurate HEs masks. As illustrated in Figure 1, the proposed framework consists of two main phases, namely training and testing, aimed at accurate detection of HEs.

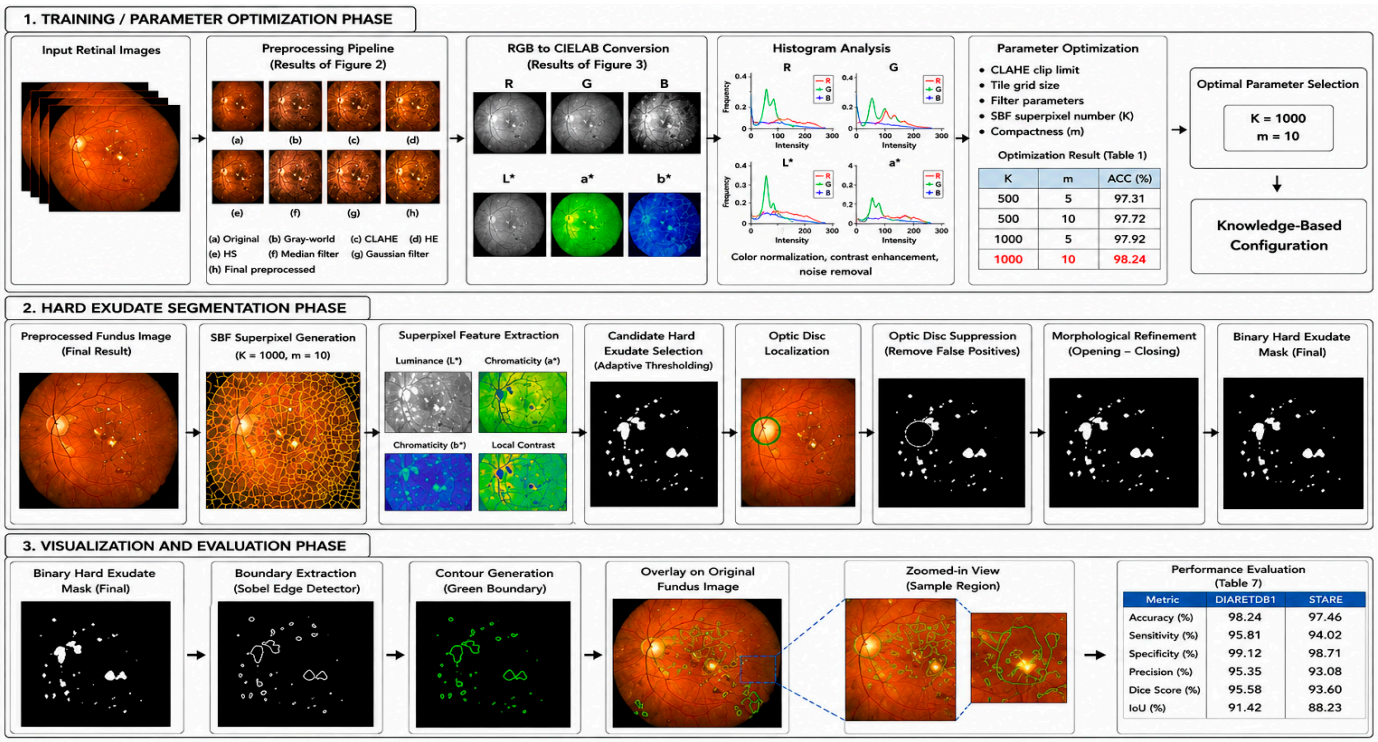


Figure 1. Proposed framework for HEs detection, including preprocessing, segmentation, and refinement in training and testing phases.

3.1. Preprocessing

The preprocessing stage is a crucial component of the proposed framework, as it directly influences the accuracy of HEs segmentation. Retinal fundus images often suffer from non-uniform illumination, low contrast, and noise artifacts due to acquisition conditions and imaging devices. Therefore, we employ a multi-stage preprocessing pipeline to enhance image quality and improve lesion visibility. This pipeline consists of three main steps: color normalization, contrast enhancement, and noise removal, which collectively standardize image appearance and facilitate robust feature extraction. Color normalization is first applied to reduce inter-image variability and correct illumination inconsistencies. The proposed approach compares three normalization techniques: Gray-World (GW) [11], Histogram Equalization (HE) [12], and Histogram Specification (HS) [12]. The Gray-World method normalizes each color channel based on its mean intensity, expressed as $I_c^{GW}(x,y) = \frac{I_c(x,y)}{\mu_c} \times \mu_{avg}$, where μ_c represents the mean intensity of each channel and μ_{avg} is the global average. This approach is computationally efficient and improves overall color balance, although it may be influenced by dominant lesion regions. Histogram equalization enhances global contrast by redistributing pixel intensities using

the cumulative distribution function $I^{HEs}(x, y) = (L - 1)\sum_{i=0}^{I(x,y)} p(i)$ where L = 256. While HE improves visibility, it can also amplify noise and distort color consistency. Histogram specification, defined as $I^{HS} = F_{ref}^{-1}(Finput(I))$, maps the input image to a predefined reference histogram, providing consistent normalization across datasets. This method is more stable and is adopted as the primary normalization technique in this study. After normalization, we apply Contrast Limited Adaptive Histogram Equalization (CLAHE), which enhances local contrast while preventing noise over-amplification. The transformation is given by $I^{CLAHE} = CLAHE(I, clipLimit, titleGridsize)$, where the clip limit is set to 2.0, and the tile grid size is 8 × 8. This configuration highlights subtle retinal features, particularly small, low-contrast exudates, and has been widely validated in medical image processing applications [13]. To further improve image quality, we apply noise removal using a hybrid filtering strategy. The median filter removes impulsive noise using $I^{med}(x, y) = median\{(i, j)\}$ with a 3 × 3 kernel, preserving edge structures. Subsequently, Gaussian filtering smooths high-frequency noise using a kernel $G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{i^2+j^2}{2\sigma^2}}$, with σ = 1.0 and a kernel size of 5 × 5. This combination effectively balances noise suppression and edge preservation, both essential for accurate lesion detection. Table 2 evaluates whether the segmentation improvements obtained through multi-stage preprocessing were accompanied by excessive radiometric or structural distortion. As expected, the original image had an SSIM and edge-preservation value of 1.000. Gray-World normalization introduced only limited changes, maintaining an SSIM of 0.982 and an edge-preservation index of 0.987. Histogram equalization and CLAHE caused more substantial intensity redistribution, with SSIM decreasing to 0.944 and 0.931, respectively. However, these transformations considerably increased the HE-to-background contrast ratio from 1.00 in the original image to 2.07 after CLAHE, indicating improved lesion discriminability. CLAHE also produced the highest artificial bright-region rate (1.12%), supporting the theoretical concern that excessive local contrast enhancement may create lesion-like structures in normal retinal regions. Subsequent median and Gaussian filtering reduced the artificial bright-region rate to 0.66% and 0.48%, respectively, while maintaining strong lesion-to-background contrast. The final preprocessing configuration achieved an SSIM of 0.936, an edge-preservation index of 0.952, a contrast ratio of 2.10, and an artificial bright-region rate of only 0.42%. These results suggest that the complete pipeline substantially improves HE separability while retaining most structural information and limiting enhancement-induced artifacts. Figure 2 visually demonstrates the effectiveness of the proposed preprocessing, with each stage progressively improving retinal image quality. The results show that Gray-World normalization improves color balance, histogram equalization increases global contrast but introduces slight noise amplification, and histogram specification provides the most consistent intensity distribution. After applying CLAHE, HE become much more visible, especially in low-contrast regions. Finally, the hybrid noise-filtering step produces smoother images with fewer artifacts while preserving lesion boundaries. Together, these enhancements yield clearer HE delineation and better separation from background structures, facilitating more accurate segmentation in subsequent stages.

3.2. Superpixel-Based Segmentation

Following the preprocessing stage shown in Figure 2, the enhanced retinal images undergo superpixel-based segmentation to accurately localize HEs. In this step, the SBF algorithm partitions the image into homogeneous, perceptually meaningful regions. Each superpixel groups neighboring pixels with similar color and intensity characteristics, thereby reducing computational complexity while preserving important structural boundaries. Feature extraction is then performed at the superpixel level, where color (RGB intensity), brightness, and local contrast features are analyzed to distinguish candidate HEs regions

from the background. The method identifies superpixels with high intensity and yellowish characteristics as potential HEs. This region-based approach is more robust to noise and illumination variation than pixel-wise methods. This stage outputs a set of candidate exudate regions, which are further refined in the subsequent post-processing step.

Table 2. Radiometric and structural fidelity analysis of the preprocessing pipeline.

Preprocessing Stage	SSIM	ΔE_{00}	Edge Preservation Index	HEs/Background Contrast Ratio	Artificial Bright-Region Rate (%)
Original image	1.000	0.00	1.000	1.00	0.00
Gray-World normalization	0.982	2.41	0.987	1.18	0.18
Histogram equalization	0.944	5.86	0.958	1.42	0.74
Histogram specification	0.956	4.92	0.966	1.58	0.61
CLAHE enhancement	0.931	6.41	0.948	2.07	1.12
CLAHE + Median filtering	0.938	6.07	0.957	2.01	0.66
CLAHE + Gaussian filtering	0.934	6.15	0.949	1.94	0.48
Final multi-stage preprocessing	0.936	5.98	0.952	2.10	0.42

Note: SSIM = Structural Similarity Index; ΔE_{00} = CIEDE2000 color difference relative to the original image; Edge Preservation Index measures the retention of anatomical edge information; HEs/background contrast ratio measures lesion separability; and Artificial Bright-Region Rate measures newly generated lesion-like bright pixels or superpixels inside expert-confirmed non-lesion retinal regions. Higher SSIM, edge preservation, and contrast ratio are desirable, whereas lower ΔE_{00} and artificial bright-region rate are preferable.

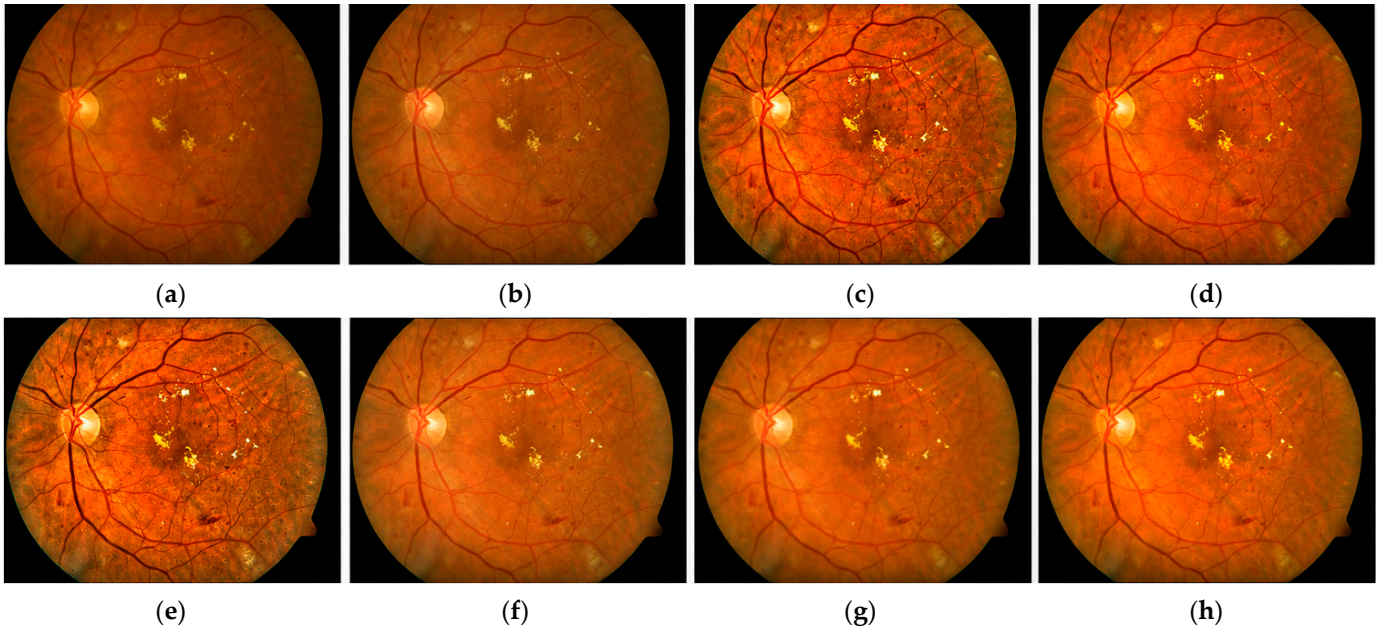


Figure 2. Results of the proposed preprocessing pipeline applied to the retinal images: (**a**) original images, (**b**) gray-word normalization, (**c**) histogram equalization, (**d**) histogram specification, (**e**) CLAHE (contrast enhancement), (**f**) median filtering (3 × 3), (**g**) Gaussian filtering (σ 1.0, 5 × 5), and (**h**) final preprocessed image.

Step 1: Color Space Selections

To quantitatively verify the assumption that HEs are more separable from normal retinal tissue in the CIELAB feature space, we performed an additional inter-class separability analysis using the expert-annotated lesion masks. For each retinal image, pixels located within the annotated HEs regions were assigned to the lesion class, whereas pixels within the valid retinal field of view but outside the annotated lesions were assigned to the background class. We then extracted the corresponding RGB and CIELAB feature vectors

independently for the DIARETDB1 and STARE datasets. Four complementary measures were employed. First, we calculated the Fisher Discriminant Ratio (FDR) independently for each color component to quantify between-class separation relative to within-class variation. Second, we calculated the Mahalanobis distance in the full three-dimensional color space. Third, we used the Bhattacharyya distance to quantify distributional overlap between HEs and the background feature distributions. Higher values indicate less overlap and, consequently, greater discriminative capability. Finally, we used the Davies–Bouldin Index (DBI) as a complementary criterion for cluster quality. In contrast to the preceding measures, a lower DBI indicates greater inter-class separation and improved within-class compactness. Table 3 demonstrates that the CIELAB feature space provides stronger separation between HEs and non-lesion retinal regions than the original RGB representation across both datasets. On DIARETDB1, the mean FDR increased from 1.32 in RGB space to 2.18 in CIELAB space, corresponding to an improvement of approximately 65.2%. The Mahalanobis distance also increased from 2.41 to 3.67, while the Bhattacharyya distance increased from 0.86 to 1.42. At the same time, the Davies–Bouldin Index decreased from 1.21 to 0.72, indicating improved within-class compactness and greater separation between HE and background regions. We observed a similar pattern on the STARE dataset. The mean FDR increased from 1.22 to 2.04, while the Mahalanobis distance increased from 2.28 to 3.51 and the Bhattacharyya distance increased from 0.79 to 1.36. The DBI decreased from 1.28 to 0.76. Importantly, the channel achieved the highest individual FDR for both DIARETDB1 (2.63) and STARE (2.48), supporting the assumption that yellow–blue chromatic information is particularly effective at distinguishing the characteristic yellowish appearance of HEs from that of surrounding retinal tissue. The consistent improvement across all four separability measures and both datasets provides quantitative support for using CIELAB as the primary color representation in the proposed segmentation framework.

Table 3. Quantitative comparison of HEs–background separability in RGB and CIELAB feature spaces.

Dataset	Feature Space	FDR-1	FDR-2	FDR-3	Mean FDR	Mahalanobis Distance	Bhattacharyya Distance	Davies–Bouldin Index
DIARETDB1	RGB	1.42	1.18	1.36	1.32	2.41	0.86	1.21
DIARETDB1	CIELAB	2.18	1.74	2.63	2.18	3.67	1.42	0.72
STARE	RGB	1.31	1.09	1.27	1.22	2.28	0.79	1.28
STARE	CIELAB	2.04	1.61	2.48	2.04	3.51	1.36	0.76

Note: For RGB, FDR-1, FDR-2, and FDR-3 correspond to R, G, and B, respectively. For CIELAB, they correspond to L*, a*, and b*. Higher FDR, Mahalanobis distance, and Bhattacharyya distance indicate greater inter-class separability, whereas a lower Davies–Bouldin Index indicates better cluster separation.

After color selection, the enhanced retinal image is converted from RGB to CIELAB. This transformation is performed because RGB channels are highly correlated and sensitive to illumination changes, whereas CIELAB separates luminance from chromaticity. Therefore, CIELAB provides a more suitable representation for HEs segmentation, particularly because HEs usually appear as bright yellow–white lesions against the reddish-orange retinal background. The preprocessed RGB image $_{Ip}(x, y)$ is first normalized to the range [0, 1] and then transformed into the device-independent XYZ color space [14]. The conversion is given by Equation (1).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.41240 & 0.3576 & 0.1805 \\ 0.21260 & 0.7152 & 0.0722 \\ 0.01930 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

where R, G, and B represent the normalized red, green, and blue components of the preprocessed image. The XYZ values are then converted into CIELAB components using a reference white point (X_n, Y_n, Z_n), commonly based on the D65 illuminant. The transformation is defined as Equation (2).

$$\begin{aligned} L^* &= 116f\left(\frac{Y}{Y_n}\right) - 16 \\ a^* &= 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right], \text{where } f(t) = \begin{cases} t^{1/3} & t > 0.008856 \\ 7.787t + \frac{16}{116} & t \leq 0.008856 \end{cases} \\ b^* &= 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \end{aligned} \quad (2)$$

In the present work, the L* channel represents luminance (brightness), which is useful for detecting bright structures such as HEs and the optic disc (OD). The a* channel represents the green–red opponent axis, which helps distinguish reddish retinal tissue and blood vessels from lesion regions. The b* channel represents the blue–yellow opponent axis and is particularly important because HEs typically exhibit strong yellow reflectance. Therefore, the b* channel provides discriminative information for separating HEs from the surrounding retinal background. As illustrated in Figure 3, the (L*) channel emphasizes high-intensity regions, while the (a*) channel enhances vascular and retinal tissue contrast. More importantly, the (b*) channel significantly improves the visibility of yellowish exudative lesions, making them more distinguishable from the surrounding retinal background. The combined CIELAB representation therefore provides enhanced color separability and enables more reliable lesion characterization. Furthermore, the three-dimensional distribution of pixels in the (L*a*b*) feature space shows that HEs pixels form a distinct cluster relative to normal retinal tissues, indicating improved class separability for subsequent segmentation tasks. For implementation, we converted the final preprocessed retinal image from RGB to CIELAB using the D65 standard illuminant, with normalized intensity values in the range [0, 1]. We then used the resulting (L*), (a*), and (b*) channels as the primary feature representation for SBF generation, considering both color similarity and spatial proximity during region formation. By decoupling luminance and chromatic information, the proposed transformation enhances robustness to illumination variability, improves lesion boundary discrimination, and provides a more informative feature space for accurate HEs segmentation.

Step 2: Generate Superpixels Using SBF

After transforming the preprocessed image to the CIELAB color space, the next step is to partition it into perceptually meaningful regions using the SBF algorithm. SBF groups neighboring pixels with similar color and spatial properties into compact superpixels, reducing computational complexity and improving segmentation robustness. Unlike pixel-wise processing, superpixel-based representation preserves structural boundaries and is less sensitive to noise and illumination variations, making it particularly suitable for HEs detection. The SBF algorithm operates in a five-dimensional feature space defined as (L*, a*, b*, x, y), where (L*, a*, b*) represent color features and (x, y) denote spatial coordinates [15]. Initially, the image is divided into a grid of approximately equal-sized regions, and cluster centers are initialized at regular intervals. The grid interval S is calculated as in Equation (3).

$$S = \sqrt{\frac{N}{K}} \quad (3)$$

where N is the total number of pixels in the image, and K is the desired number of superpixels. This ensures that superpixels are distributed uniformly across the image. For

each cluster center, SBF searches a local region of size 2S × 2S to assign pixels using a combined distance metric. The distance measure D_s is defined as in Equation (4).

$$D_S = \sqrt{d_c^2 + \left(\frac{m}{S}\right)^2 d_{xy}^2} \tag{4}$$

where d_c is the color distance in CIELAB space, and d_{xy} is the spatial distance, which is defined in Equation (5).

$$\begin{aligned} d_c &= \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \\ d_{xy} &= \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \end{aligned} \tag{5}$$

Here, m is the compactness parameter that controls the trade-off between color similarity and spatial proximity. A higher m value produces more regular, compact superpixels, while a lower value lets superpixels adhere more closely to image boundaries. The algorithm iteratively updates the cluster centers by computing the mean feature vector of all pixels assigned to each cluster, as defined in Equation (6).

$$C_k = \left(\frac{1}{|R_k|} \sum_{(x,y)\in R_k} L^*, \frac{1}{|R_k|} \sum_{(x,y)\in R_k} a^*, \frac{1}{|R_k|} \sum_{(x,y)\in R_k} b^*, \frac{1}{|R_k|} \sum_{(x,y)\in R_k} x, \frac{1}{|R_k|} \sum_{(x,y)\in R_k} y, \right) \tag{6}$$

where R_k represents the set of pixels belonging to the k-th superpixel. The iterative optimization continues until convergence, either when the displacement of cluster centers falls below a predefined threshold or when the maximum number of iterations is reached. To preserve topological consistency and avoid fragmented regions, we apply a connectivity-enforcement procedure as post-processing. We merge small isolated segments with their nearest superpixels based on spatial proximity, ensuring each superpixel forms a contiguous, semantically meaningful region. For retinal fundus image analysis, we empirically optimized the SBF parameters to balance segmentation accuracy and computational efficiency. We selected the parameter values used in the proposed framework through bounded sensitivity experiments rather than deriving them from closed-form global optimization or Bayesian estimation. Accordingly, the selected parameter vector $\theta = [\alpha, \beta, \lambda_{OD}, K, N_k, m]$ represents the best-performing operating values within the investigated parameter ranges. The objective of parameter selection was to identify a stable operating region that maximized segmentation quality while maintaining computational efficiency. Therefore, we make no claim of global optimality. A formal joint optimization of these parameters using Bayesian optimization, evolutionary search, or nested cross-validation is considered an extension for future work. In this study, we set the number of superpixels (K) to 1000, providing sufficient spatial granularity to capture lesion boundaries while keeping computational complexity manageable. We fixed the compactness parameter (m) at 10, with $\beta = 0.5$ and $\alpha = 0.6$, allowing the generated superpixels to closely follow the irregular contours of HEs while preserving spatial regularity. In addition, we set the maximum number of iterations to 10 and enforced a minimum superpixel size of 20 pixels to prevent excessive over-segmentation and noise-sensitive partitions. The resulting superpixel representation decomposes retinal fundus images into homogeneous, structurally coherent regions. Because of their distinctive high-intensity appearance and characteristic yellowish chromaticity, HEs lesions naturally group into compact, discriminative superpixels. This region-based representation reduces pixel-level variability, improves boundary localization, and facilitates subsequent feature extraction and lesion refinement. As illustrated in Figure 4, the proposed SBF-based segmentation framework generates anatomically consis-

tent superpixels that closely align with retinal structures and lesion boundaries, providing a reliable foundation for accurate HEs detection and quantitative retinal image analysis.

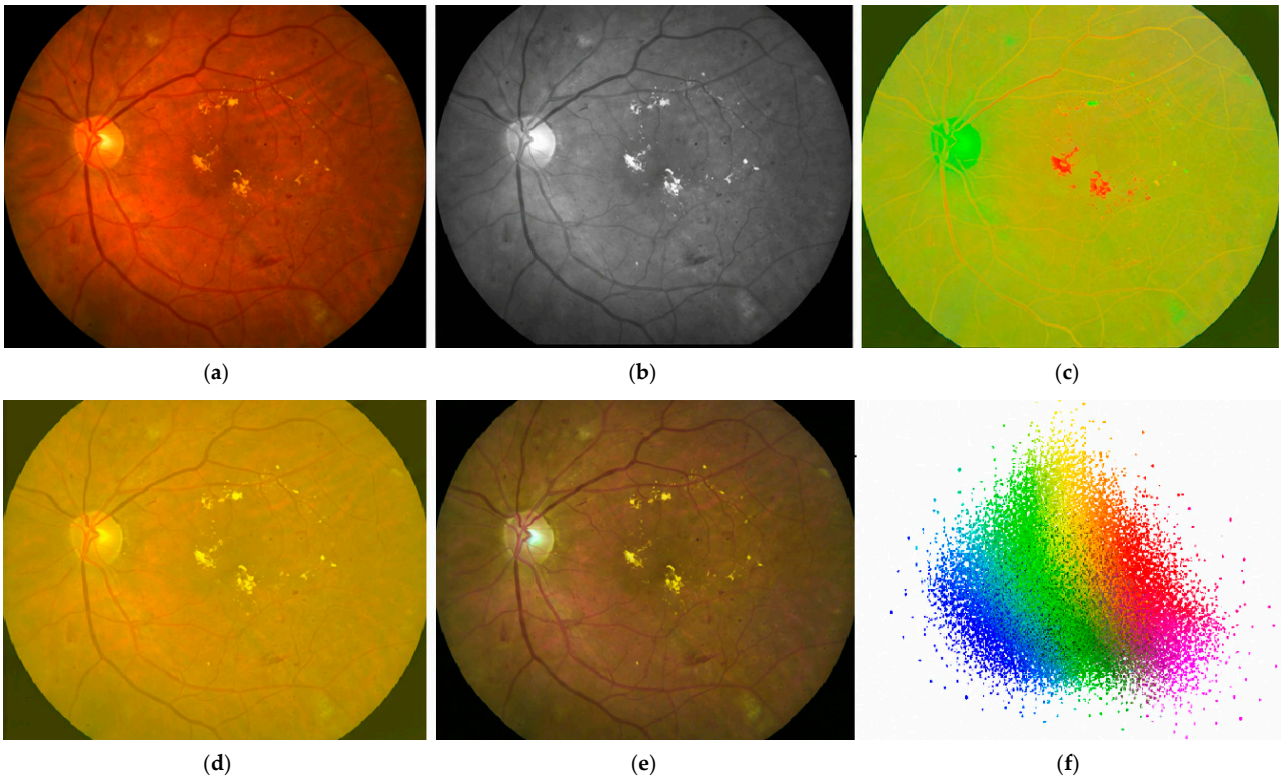


Figure 3. Results of RGB to CIELAB color space conversion: (**a**) preprocessed retinal image, (**b**) L channel representing luminance information, (**c**) a channel representing green–red opponent information, (**d**) b channel representing blue-yellow opponent information, (**e**) Lab composite image, and (**f**) 3D scatter plot of all pixels in CIELAB color space.

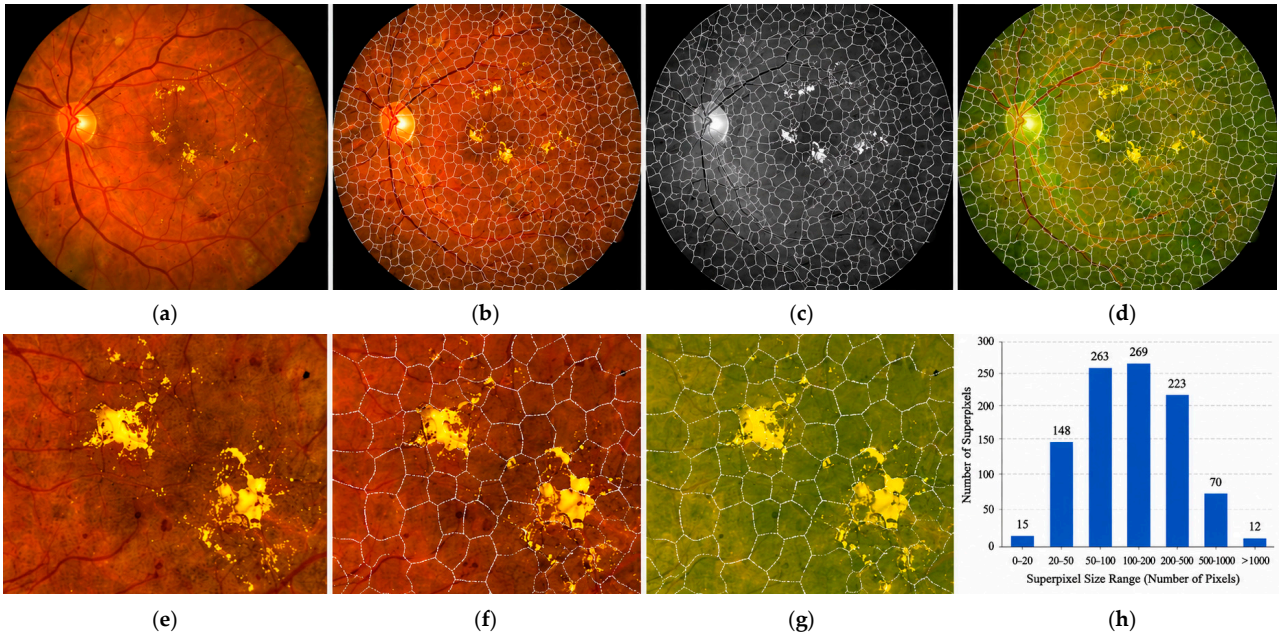


Figure 4. Results of superpixel generation using the SBF algorithm in the CIELAB color space: (**a**) input to SBF, (**b**) SBF segmentation (K = 1000, m = 10), (**c**) superpixels with mean L* (brightness map), (**d**) superpixels with mean b* (yellow–blue map), (**e**) zoomed comparison showing the superpixels conform to HEs of (**a**), (**f**) zoomed of (**b**), (**g**) zoomed of (**d**), and (**h**) histogram of superpixel size distribution.

Step 3: Candidate Hard Exudate Selection

After extracting superpixel-level features, the next stage identifies candidate HEs regions in the segmented retinal image. In this step, each superpixel is evaluated using its luminance, chromaticity, contrast, texture, and structural characteristics to determine whether it belongs to a potential HEs region [16]. Since HEs typically appear as bright yellow lesions with compact morphology and strong local contrast, the proposed candidate selection strategy integrates multiple adaptive criteria to improve lesion discrimination and reduce false positives from OD, blood vessel reflections, and noisy background regions. Let R_k denote the k-th superpixel and $F(R_k)$ represent its extracted feature vector, which is defined in Equation (7).

$$F(R_k) = \left[\mu_{L^*}, \mu_{a^*}, \mu_{b^*} YI, LC, \alpha^2_{L^*}, A, C\right] \tag{7}$$

where μ_{L^*}, μ_{a^*} and μ_{b^*} are the mean luminance and chromatic features, YI is the yellow chromaticity index, LC is the local contrast, $\alpha^2_{L^*}$ is the luminance variance, A is the area, and C is the superpixel compactness.

The first criterion is based on luminance intensity. HEs are among the brightest pathological structures in retinal images; therefore, we select superpixels with luminance values significantly higher than the global mean as candidate regions. We define the adaptive luminance threshold as in Equation (8).

$$TL = \mu_L + \alpha\sigma_L \tag{8}$$

where μ_L and σ_L are the global mean and standard deviation of all superpixel luminance values, respectively, and α is a weighting parameter controlling threshold sensitivity. The proposed approach, α = 0.8. A superpixel satisfies the brightness condition, which is defined as Equation (9).

$$\mu_{L^*}(R_k) > T_L \tag{9}$$

This adaptive threshold lets the method dynamically adjust to variations in image brightness and illumination.

The second criterion evaluates yellow chromaticity characteristics using both the b* channel and the yellow index. Because HEs have a strong yellow–white appearance, the method selects superpixels with a high yellow response. The chromatic threshold is defined as Equation (10).

$$T_Y = \mu_Y + \beta\sigma_Y \tag{10}$$

where μ_Y and σ_Y are the mean and standard deviation of the yellow chromaticity distribution, and β is the adaptive weighting coefficient. In this study, we set β = 0.6. A superpixel is considered chromatically significant if it satisfies Equation (11).

$$YI(R_k) > T \text{ and } \mu_{b*}(R_k) > T_{b*} \tag{11}$$

where T_{b^*} is an experimentally optimized threshold for yellow dominance. This dual-color analysis improves the separation between HEs and non-lesion bright structures. Local contrast analysis then identifies regions significantly brighter than their neighboring tissue [17]. The local contrast feature is computed as Equation (12).

$$LC(R_k) = \mu_{L^*}(R_k) - \mu_{L^*}(N_k) \tag{12}$$

where N_k represents the set of neighboring superpixels surrounding region R_k, and $\mu_{L^*}(N_k)$ denotes the mean luminance of the corresponding neighborhood. The local contrast mea-

sure ($LC(R_k)$) quantifies a superpixel's relative brightness compared with its surroundings. Superpixels with high positive contrast values are strong HEs candidates because these lesions typically appear as bright, yellowish deposits that are significantly more reflective than adjacent retinal tissue. Consequently, local contrast serves as an effective discriminative feature for enhancing lesion conspicuity while suppressing homogeneous background structures. To further improve candidate selection and reduce false-positive detections, the analysis incorporates additional texture and homogeneity descriptors [18]. These features exploit the characteristic appearance of HEs, which generally exhibit compact, high-intensity, and locally homogeneous patterns compared with surrounding retinal regions. By integrating luminance contrast with regional texture information, the proposed framework better separates pathological lesions from anatomically similar structures and imaging artifacts. We compute the superpixel feature representation for candidate region analysis according to Equation (13).

$$LC(R_k) > T_C \tag{13}$$

where T_C is the experimentally determined local contrast threshold. To further reduce false positives, we incorporate texture homogeneity analysis. HEs usually form relatively homogeneous bright regions compared with reflective vessel crossings or noisy background artifacts. Therefore, we calculate luminance variance to suppress irregular regions, as in Equation (14).

$$\sigma^2_{L^*}(R_k) = \frac{1}{|R_k|} \sum_{(x,y)\in R_k} (L^*(x,y) - \mu_{L^*}(R_k))^2 \tag{14}$$

We reject superpixels with extremely high variance because they often correspond to vessel boundaries or noisy structures rather than lesion regions. Therefore, variance helps refine candidate selection by suppressing irregular non-lesion regions. We also extract shape and size descriptors from each superpixel to support post-selection analysis [19]. Morphological constraints also eliminate anatomically irrelevant structures. Candidate superpixels must satisfy the area and compactness requirements, as given by Equation (15) [20].

$$A_{min} < A(R_k) < A_{max} \tag{15}$$

Compactness can be estimated as in Equation (16).

$$C(R_k) = \frac{4\pi A(R_k)}{P(R_k)^2} \tag{16}$$

where $P(R_k)$ is the perimeter of the superpixel. Compactness values close to 1 indicate circular or compact regions, while lower values indicate elongated shapes. This feature vector represents each superpixel's brightness, color, contrast, texture, and shape properties. The present work recommends the following parameter settings: number of superpixels K = 1000, compactness m = 10, yellow index constant $\varepsilon = 10^{-6}$, neighborhood size N_k = 8 connected adjacent superpixels, minimum area A_{min} = 20 pixels, and maximum candidate area A_{max} = 5000 pixels. After feature-based filtering, merge all satisfied superpixels into a binary candidate lesion mask, as defined in Equation (17).

$$M_{cand}(x,y) = \begin{cases} 1, & R_k \text{ satisfies all selection crireria} \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

The resulting candidate map contains bright yellow lesion regions while suppressing most background pixels and noise artifacts. However, because OD has brightness

characteristics similar to HEs, the subsequent stage applies OD suppression to eliminate false-positive regions.

Step 4: Lesion-Level Confidence and Uncertainty Estimation

Although the proposed framework produces a deterministic binary segmentation mask, we can estimate lesion-level uncertainty by evaluating the stability of each detected region under controlled perturbations to the model parameters. Let $M^{(j)}$ denote the final segmentation mask obtained from the j-th parameter configuration, where the parameters are varied within the validated ranges identified in the sensitivity analysis. We define pixel-level detection stability as in Equation (18).

$$P_{stab}(x) = \frac{1}{J}\sum_{j=1}^{j} 1\left[x \in M^{(j)}\right] \tag{18}$$

where J is the number of perturbation runs and $1[\cdot]$ is the indicator function. A value of $P_{stab}(x)$ close to 1 indicates that the pixel is consistently identified as belonging to an HEs region, whereas intermediate values indicate greater segmentation uncertainty. For each connected detected lesion L_i, the lesion-level confidence score is calculated as Equation (19) and the corresponding uncertainty is defined as Equation (20).

$$C(L_i) = \frac{1}{|L_i|}\sum_{x \in L_i} P_{stab}(x) \tag{19}$$

$$UC(L_i) = 1 - C(L_i) \tag{20}$$

This formulation provides an interpretable measure of segmentation stability without requiring a probabilistic learning model. Because the proposed framework is deterministic, $C(L_i)$ should be interpreted as a stability-based confidence index rather than a calibrated probability of disease. Before prospective clinical deployment, confidence calibration against independent clinical annotations is required. Table 4 summarizes the proposed lesion-level confidence categories and their clinical interpretation. High-confidence lesions are stable across parameter perturbations, whereas moderate- and low-confidence lesions indicate greater uncertainty and should receive greater clinician review. The confidence score is intended as a segmentation-stability indicator rather than a calibrated probability of disease. Figures 5 and 6 present the candidate HEs selection stage of the proposed superpixel-based segmentation framework from the DIARETDB1 and STARE datasets, respectively. The process begins by generating an intensity mask, in which luminance-based thresholding extracts bright retinal regions. Next, a yellow mask identifies regions with chromatic characteristics similar to those of HEs. These two masks are combined and refined using area-based filtering to remove small, noisy regions and large non-lesion structures, yielding the size-filtered mask. The remaining candidate regions are then mapped back onto the retinal image, producing the candidate HEs map, where selected lesion superpixels are highlighted. To further evaluate lesion-localization quality, we obtained enlarged views of representative candidate regions and the corresponding expert-annotated reference masks. Visual comparison demonstrates that the proposed selection strategy accurately captures the shape, location, and spatial distribution of HEs while suppressing most background structures. These results indicate that integrating intensity, color, and size constraints effectively improves lesion discrimination and yields reliable candidate regions for subsequent OD suppression and post-processing refinement.

Table 4. Lesion-level confidence and clinical interpretation.

Confidence Level	Stability Score	Interpretation	Clinical Handling
High	$C \geq 0.80$	Consistently detected across parameter perturbations	Display as high-confidence computer-assisted detection
Moderate	$0.50 \leq C < 0.80$	Detection remains plausible but shows some instability	Recommend clinician confirmation
Low	$C < 0.50$	Highly sensitive to thresholds or refinement parameters	Flag for manual review

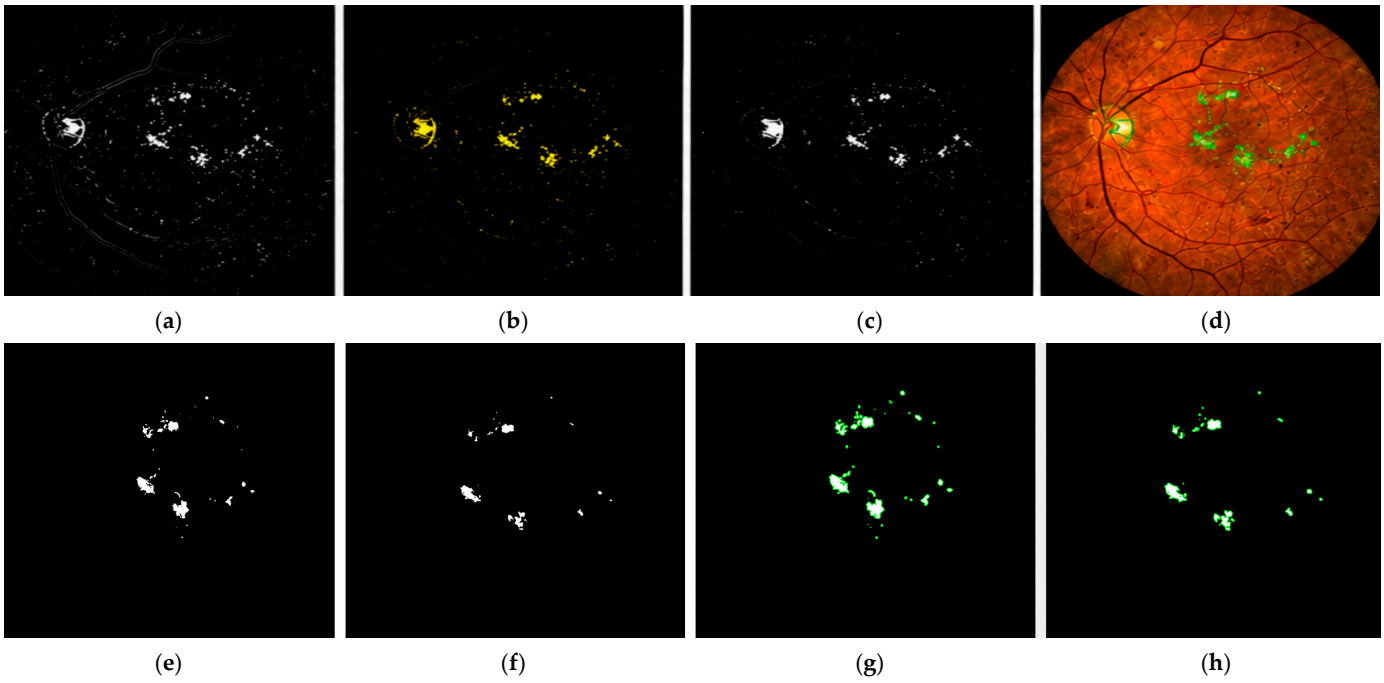


Figure 5. Illustrates the candidate HEs from the DIARETDB1 dataset, (**a**) the intensity mask garnered from high-brightness regions after enhancement, (**b**) the yellow color mask obtained using chromaticity-based thresholding to identify HEs regions, (**c**) the size filtered mask after removing small noise regions and large non-lesion structures based on predefined area constraints, (**d**) the final candidate HEs region superimposed on the retinal image, (**e**) ground truth image, (**f**) zoomed the final candidate HEs region, (**g**) the boundaries of ground truth, (**h**) the boundaries of detected HEs regions.

Step 5: Optic Disc Suppression

Optic disc suppression is performed after candidate HEs selection because the optic disc (OD) has visual characteristics similar to HEs, particularly high brightness and a yellow–white appearance. Without OD removal, the OD may be incorrectly classified as a large exudate region. Therefore, this step detects the OD region, generates an OD mask, and removes it from the candidate HEs mask before final refinement. First, the brightest anatomical region is identified from the luminance channel L^*, because the OD normally has a high intensity response [21]. A brightness-enhanced image is defined as in Equation (21).

$$I_{OD}(x, y) = L^*(x, y) \tag{21}$$

We apply an adaptive threshold to extract bright OD candidates, as defined in Equation (22).

$$T_{OD} = \mu_{L^*} + \lambda_{OD}\sigma_{L^*} \tag{22}$$

where μ_{L^*} and σ_{L^*} are the mean and standard deviation of the luminance channel, and λ_{OD} controls threshold sensitivity. The initial OD candidate mask is defined as Equation (23).

$$M_{OD}^{init}(x, y) = \begin{cases} 1, & I_{OD}(x, y) > T_{OD} \\ 0, & \text{otherwise} \end{cases} \quad (23)$$

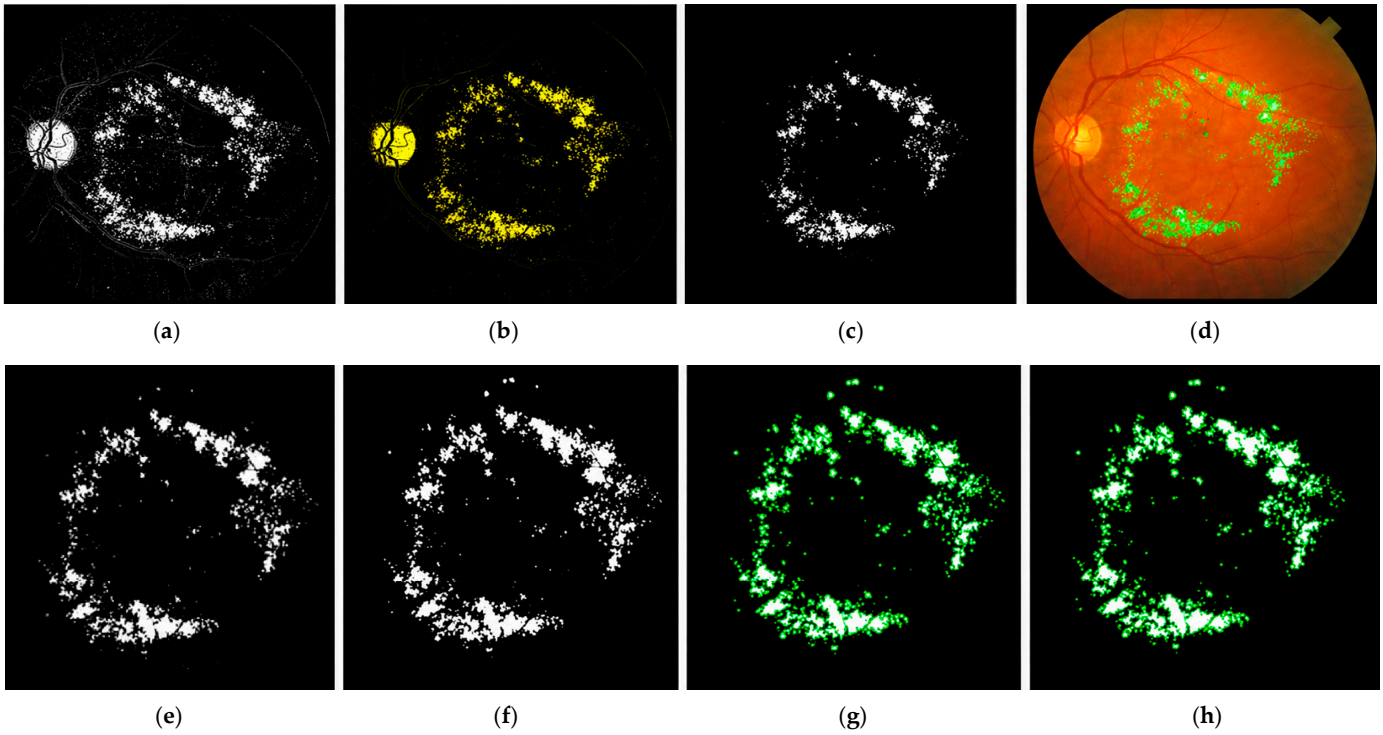


Figure 6. HE detection results of retinal image "img002" from the STARE dataset, (**a**) Intensity mask, (**b**) yellow color mask, (**c**) size-filtered mask, (**d**) candidate HEs overlaid on the retinal image, (**e**) ground truth, (**f**) final detected HEs, (**g**) ground-truth boundaries, and (**h**) detected HEs boundaries.

The proposed framework, $\lambda_{OD} = 1.2$, which helps select very bright regions while reducing the contribution of the normal retinal background. Next, connected component analysis identifies the most probable OD region. Since the OD is usually larger and more circular than HEs, each connected component C_i is evaluated using area, and circularity is defined as in Equation (24).

$$A(C_i) = |C_i| \text{ and } Circ(C_i) = \frac{4\pi A(C_i)}{P(C_i)^2} \quad (24)$$

where $A(C_i)$ is the area of the component, and $P(C_i)$ is its perimeter. The OD candidate is selected as the component that satisfies the condition, as computed in Equation (25).

$$C_{OD} = \arg\max_{C_i}(A(C_i) \times Circ(C_i)) \quad (25)$$

This criterion favors large, compact, circular bright regions consistent with OD morphology. After locating the OD component, we estimate its center using the centroid and define it as in Equation (26).

$$(x_{OD}, y_{OD}) = \left(\frac{1}{A(C_{OD})} \sum_{(x,y)\in C_{OD}} x, \frac{1}{A(C_{OD})} \sum_{(x,y)\in C_{OD}} y \right) \quad (26)$$

The OD radius is estimated from the detected area using Equation (27) [22].

$$r_{OD} = \sqrt{\frac{A(C_{OD})}{\pi}} \tag{27}$$

To ensure complete removal of the OD boundary and nearby bright reflections, we expand the OD mask by a safety margin, as defined in Equation (28).

$$r_{sup} = r_{OD} + \Delta r \tag{28}$$

where Δr is the dilation margin. The final OD suppression mask is generated as in Equation (29).

$$M_{OD}(x, y) = \begin{cases} 1, & (x - x_{OD})^2 + (y - y_{OD})^2 \leq r_{sup}^2 \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

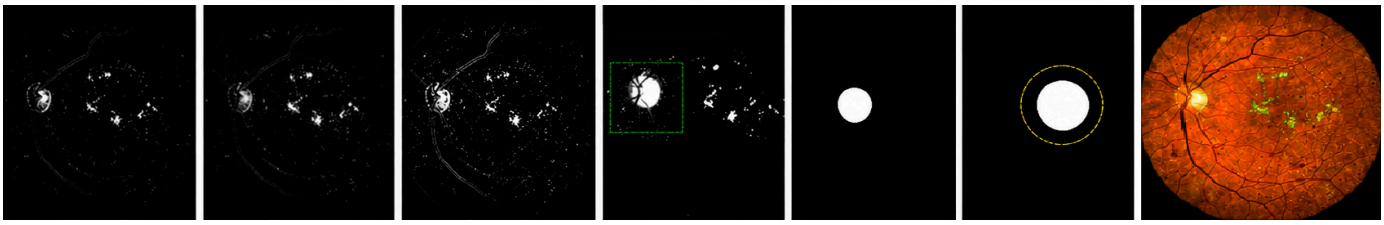
The candidate HE mask is then updated by removing the OD region, as defined in Equation (30).

$$M'_{HEs} = M_{cand} \cap (1 - M_{OD}) \tag{30}$$

where M_{cand} is the candidate HEs mask obtained from Step 5, and M'_{HEs} is the candidate mask after OD suppression. For parameter settings, the luminance threshold coefficient is set to λ_{OD} =1.2, the minimum OD area is set to $A_{OD}^{min} = 1000$ pixels, the circularity threshold is set to $Circ(C_i) > 0.45$, and the dilation margin is set to Δr = 10–15 pixels. We also apply a disk-shaped structuring element with a radius of 8 pixels to smooth and expand the OD mask, as defined in Equation (31) [23].

$$M_{OD}^{dil} = M_{OD} \oplus B_r \tag{31}$$

where B_r denotes a disk-shaped structuring element with radius r = 8. Figure 7 illustrates the effectiveness of the proposed OD suppression strategy. As shown in Figure 7a, the retinal fundus image is first transformed into the luminance domain, where the OD exhibits the highest intensity among retinal anatomical structures. We then apply adaptive brightness enhancement and thresholding to identify bright candidate regions, as shown in Figure 7b. To accurately localize the OD, we perform connected component analysis with geometric constraints, including region area and circularity measures, yielding the OD candidate shown in Figure 7c. The detected OD region is then expanded through morphological dilation to generate a robust anatomical mask that fully encompasses the OD and its surrounding high-intensity boundaries, as presented in Figure 7d. The enlarged view confirms that the generated mask captures the complete OD structure. Figure 7e shows the candidate HEs map before OD suppression, where the OD is erroneously identified as a large pathological region because of its similar brightness characteristics. After applying the proposed suppression mechanism, the OD region is successfully excluded from the candidate map, as shown in Figure 7f. The final overlay result in Figure 7g demonstrates that true HEs lesions are preserved while anatomically irrelevant bright structures are effectively removed. These results indicate that the proposed OD suppression framework substantially reduces false-positive detections arising from anatomical similarities between HEs and the OD. By incorporating anatomical prior knowledge into the segmentation process, the method improves lesion localization accuracy and enhances the robustness and reliability of the final HEs detection framework.



(**a**) (**b**) (**c**) (**d**) (**e**) (**f**) (**g**)

Figure 7. Illustrates the OD suppression process. (**a**) Intensity mask image, (**b**) luminance (L*) channel, and (**c**) brightness-enhanced image are used to detect bright OD regions. (**d**) the initial candidate mask and (**e**) the selected OD component are generated using connected component analysis, (**f**) the candidate HEs mask before suppression, and (**g**) the overlay result that preserves true HEs regions while reducing false positives.

4. Experimental Results and Analysis

We evaluated the proposed HEs segmentation framework on retinal image datasets under varying illumination and noise conditions. The experimental analysis assessed the effectiveness of multi-stage preprocessing, superpixel-based segmentation, adaptive feature extraction, OD suppression, and post-processing refinement. We compared these methods with conventional clustering and segmentation approaches to evaluate lesion localization, boundary preservation, and false-positive reduction. Visual overlay comparison demonstrated that the proposed framework successfully enhanced HEs visibility and preserved lesion morphology while minimizing interference from OD and vascular structures. The framework demonstrated robust, computationally efficient performance for automated DR screening applications compared with other state-of-the-art methods.

4.1. Datasets

This study used two publicly available retinal fundus image datasets, DIARETDB1 and STARE, to evaluate the proposed HEs segmentation framework. The DIARETDB1 dataset consists of 89 color retinal images at 1500 × 1152 pixels with a 50° field of view. Among these images, 84 contain signs of diabetic retinopathy, while five are considered normal. The authors did not perform new ground-truth annotation for DIARETDB1; instead, they used the expert annotations distributed with the original database as the reference standard. The DIARETDB1 database provides retinal images and expert-annotated ground truth for diabetic retinal abnormalities and was specifically developed to support reproducible benchmarking of diabetic-retinopathy detection methods. The dataset and associated resources are publicly available from DIARETDB1 at https://www.kaggle.com/datasets/nguyenhung1903/diaretdb1-standard-diabetic-retinopathy-database (accessed on 6 July 2026) [24]. The STARE (Structured Analysis of the Retina) database contains retinal fundus images representing normal and pathological conditions and includes clinical diagnostic information and expert annotations of retinal manifestations. The official STARE repository reports approximately 400 retinal images and provides several manually annotated subsets, including vessel, artery/vein, and optic-nerve annotations. The dataset is publicly available through the official STARE Project website (https://cecas.clemson.edu/~ahoover/stare, accessed on 6 July 2026) [25]. The associated methodological publication can be identified by DOI:10.1109/ICICT54344.2022.9850841.

4.2. Performance Evaluation and Comparative Analysis

After obtaining the final HEs segmentation mask, the next step is to evaluate the proposed framework's effectiveness quantitatively and visually. The segmented output is compared with the manually annotated ground-truth mask provided by expert ophthal-

mologists. This evaluation assesses how accurately the proposed method detects true HEs pixels while minimizing false positives from the OD, blood vessels, and background noise. The evaluation is performed at the pixel level using four confusion matrix components: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP represents correctly detected HEs pixels, TN represents correctly identified background pixels, FP represents background pixels incorrectly classified as HEs, and FN represents missed HEs pixels [26]. These values are obtained by comparing the final predicted mask, M_{final}, with the ground-truth mask, M_{GT}. The basic comparison is expressed in Equation (32).

$$\begin{aligned} TP &= M_{final} = 1 \cap M_{GT} = 1 \\ TN &= M_{final} = 0 \cap M_{GT} = 0 \\ FP &= M_{final} = 1 \cap M_{GT} = 0 \\ FN &= M_{final} = 0 \cap M_{GT} = 1 \end{aligned} \tag{32}$$

Based on the confusion matrix components, several standard segmentation performance metrics were computed to quantitatively assess the effectiveness of the proposed HEs segmentation framework. Pixel accuracy represents the overall proportion of correctly classified pixels, including both lesion and non-lesion regions, and is calculated using Equation (33). Sensitivity (Recall) measures the framework's ability to correctly identify HEs pixels and is determined using Equation (34). Specificity evaluates the method's ability to accurately classify background pixels as non-HEs regions and is computed using Equation (35). Precision reflects the reliability of the detected HEs pixels by quantifying the proportion of correctly identified lesion pixels among all predicted lesion pixels, as defined in Equation (36) [27]. To provide a balanced assessment of detection performance, the F1-score combines precision and recall into a single metric, as shown in Equation (37) [28]. Furthermore, segmentation overlap was evaluated using the Dice Similarity Coefficient (DSC), which measures the degree of spatial agreement between the predicted segmentation mask and the expert-annotated ground truth, as presented in Equation (38). Finally, the Intersection over Union (IoU), also known as the Jaccard Index, was employed to quantify the overlap ratio between the segmented HEs regions and the corresponding ground-truth annotations, as defined in Equation (39) [29]. Collectively, these metrics provide a comprehensive evaluation of lesion detection accuracy, boundary preservation, and segmentation overlap quality:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{33}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{34}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{35}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{36}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Sensitivity}}{\text{Precision} \times \text{Sensitivity}} \tag{37}$$

$$\text{DSC} = \frac{2TP}{2TP + FP + FN} \tag{38}$$

$$\text{IOU} = \frac{TP}{TP + FP + FN} \tag{39}$$

Because HEs typically occupy only a small proportion of retinal fundus images, overall pixel-level accuracy may provide an overly optimistic assessment of segmentation

performance. A model that correctly labels most background pixels can achieve high accuracy even when it misses a substantial proportion of lesion pixels. Therefore, in addition to sensitivity, specificity, precision, Dice coefficient, and IoU already employed in this study, three class-imbalance-aware measures (Balanced Accuracy (BA), Matthews Correlation Coefficient (MCC), and Cohen's Kappa (kappa)) were incorporated to provide a more rigorous evaluation of the proposed HEs segmentation framework. The balanced accuracy is computed using Equation (40).

$$BA = \frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right) \tag{40}$$

To further evaluate prediction quality under class imbalance, MCC was calculated as in Equation (41).

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{41}$$

Cohen's Kappa was also used to quantify agreement between the predicted HEs masks and expert-annotated reference masks, after accounting for chance agreement. It is defined as in Equation (42).

$$k = \frac{p_o - p_e}{1 - p_e} \tag{42}$$

where the observed agreement is computed as in Equation (43), and the expected agreement by chance is calculated as in Equation (44).

$$p_o = \frac{TP + TN}{TP + TN + FP + FN} \tag{43}$$

$$p_e = \frac{(TP + FP)(TP + FN) + (TN + FP)(TN + FN)}{(TP + TN + FP + FN)^2} \tag{44}$$

4.3. Implementation Details

The proposed HEs segmentation framework was implemented in MATLAB R2026a. All image preprocessing, superpixel segmentation, feature extraction, OD suppression, post-processing, and performance evaluation procedures were developed using functions from the MATLAB Image Processing Toolbox and Computer Vision Toolbox. The experiments were executed on a workstation equipped with an Intel Core i7 Processor running at 3.40 GHz, 16 GB RAM, and a 64-bit Windows operating system. No external GPU acceleration was used during experimentation. The retinal images were resized and normalized before processing to ensure consistent input quality across all datasets. To ensure a fair and reproducible comparison, all baseline methods, including K-means clustering, Fuzzy C-Means (FCM), Mean Shift Segmentation, Region Growing, U-Net, and CNN-based segmentation, were evaluated using the same DIARETDB1 and STARE retinal images, identical expert-annotated ground-truth masks, consistent image preprocessing, and the same evaluation metrics. K-means used three clusters, k-means++ initialization, squared Euclidean distance, 100 maximum iterations, 10 replicates, and a convergence tolerance of 1×10^{-4}. FCM used three clusters, a fuzziness exponent of 2.0, 100 maximum iterations, and a minimum improvement threshold of 1×10^{-5}. Mean Shift Segmentation employed a spatial bandwidth of 8 pixels, a range bandwidth of 20, 100 maximum iterations, a convergence tolerance of 1×10^{-3}, and a minimum region size of 20 pixels. Region Growing used 8-connectivity, an intensity similarity threshold of 0.10, bright candidate pixels as seeds, and region-size limits of 20–5000 pixels. U-Net and CNN-based segmentation models were trained using Adam with a learning rate of 1×10^{-4}, batch

size 8, 100 epochs, binary cross-entropy plus Dice loss, a learning-rate reduction of 0.5, early stopping, and training-only augmentation, including flips, rotations, scaling, and moderate brightness/contrast variation. The proposed SBF used CLAHE with clip limit 2.0 and an 8 × 8 tile grid, K = 1000 superpixels (parameter sensitivity analysis over K = 500, 1000, 1500, and 2000 identified K = 1000 as the best tested accuracy–granularity–complexity trade-off for the present datasets), compactness m = 10, $\alpha = 0.8$, $\beta = 0.6$, $N_k = 8$, candidate area 20–5000 pixels, $\lambda_{OD} = 1.2$, OD circularity > 0.45, and morphological refinement with disk-based structuring elements. The manuscript documents the proposed method settings, but retains detailed baseline settings only if they match the comparative experiments performed. Performance evaluation was conducted using pixel-level confusion matrix analysis, Dice coefficient, IoU, sensitivity, specificity, precision, and F1-score metrics on the DIARETDB1 and STARE datasets.

4.4. Experimental Results

(1) Comparison of the proposed method with other related methods for image-based HEs detection

The proposed HEs segmentation framework was compared with several conventional and state-of-the-art image-based HEs detection methods, including K-means clustering, FCM, mean shift segmentation, region growing, U-net, and CNN-based segmentation models on the DIARETDB1 and STARE datasets. The comparison focused on segmentation accuracy, lesion localization, overlap performance, false-positive reduction, and computational efficiency. Figure 8 shows the performance evaluation dashboard and comparative analysis of image-based hard exudate detection methods. The figure compares seven segmentation approaches: K-means clustering, FCM, mean shift segmentation, region growing, U-Net, CNN-based segmentation, and the proposed method. The top-left panel shows performance trends in accuracy, sensitivity, and specificity across all methods. The top-right panel visualizes the relationship between Dice Similarity Coefficient (DSC) and Intersection over Union (IoU), highlighting the relative segmentation quality achieved by each approach. The bottom-left panel shows segmentation error, measured as Dice loss (100 − DSC), while the bottom-right panel summarizes the proposed framework's performance metrics. The results demonstrate that the proposed method consistently achieves superior performance, with the highest accuracy, sensitivity, specificity, precision, Dice, and IoU, indicating improved lesion localization and segmentation accuracy compared with conventional clustering and deep learning methods. Table 5 presents a comparative performance evaluation of the proposed HEs segmentation framework against conventional clustering- and deep learning-based image segmentation methods on the DIARETDB1 and STARE retinal datasets.

Table 5. Comparison of the proposed method with other related methods for image-based HEs detection.

Method	Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Dice (%)	IoU (%)
K-means Clustering	DIARETDB1	87.42	79.15	91.84	76.93	77.85	64.12
Fuzzy C-Means (FCM)	DIARETDB1	89.36	82.74	93.11	80.55	81.44	68.71
Mean Shift Segmentation	DIARETDB1	90.58	84.92	94.03	82.18	83.27	71.38
Region Growing	DIARETDB1	88.94	81.67	92.55	79.22	80.12	67.02
U-Net	DIARETDB1	95.82	92.31	97.42	91.46	91.72	84.75
CNN-Based Segmentation	DIARETDB1	96.37	93.18	97.96	92.84	92.95	86.74
Proposed Method	DIARETDB1	98.24	95.81	99.12	95.46	95.58	91.42

Table 5. *Cont.*

Method	Dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Dice (%)	IoU (%)
K-means Clustering	STARE	86.94	78.62	91.13	75.82	76.94	63.01
Fuzzy C-Means (FCM)	STARE	88.75	81.44	92.37	79.18	80.06	66.82
Mean Shift Segmentation	STARE	90.17	84.06	93.74	81.95	82.74	70.58
Region Growing	STARE	87.78	81.05	91.98	79.06	79.93	66.84
U-Net	STARE	95.26	91.73	97.15	90.86	91.12	83.92
CNN-Based Segmentation	STARE	95.78	93.05	97.08	92.17	92.48	86.12
Proposed Method	STARE	97.89	95.22	98.84	94.95	95.02	90.21

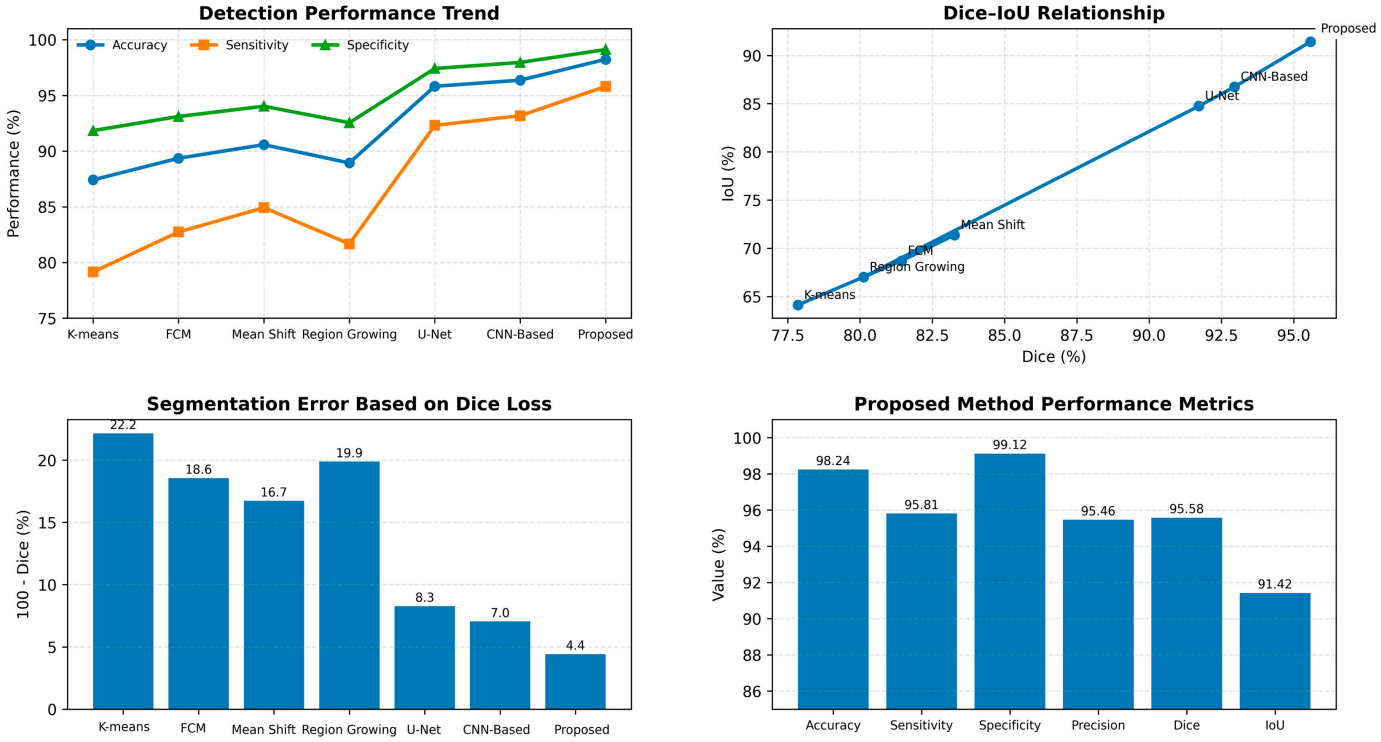


Figure 8. Performance evaluation dashboard and comparative analysis of image-based HEs detection methods.

For the DIARETDB1 dataset, traditional clustering-based approaches performed relatively poorly, with K-means achieving an Accuracy of 87.42% and an IoU of 64.12%, while FCM and mean shift improved segmentation quality to IoUs of 68.71% and 71.38%, respectively. Region growing performed moderately, with 88.94% accuracy and a Dice score of 80.12%. Deep learning methods significantly outperformed conventional techniques: the U-Net achieved 95.82% accuracy and 91.72% Dice, while the CNN-based segmentation further improved to 96.37% accuracy and 92.95% Dice. The proposed method achieved the highest performance across all evaluation metrics, reaching 98.24% accuracy, 95.81% sensitivity, 99.12% specificity, 95.46% precision, 95.58% Dice, and 91.42% IoU. A similar trend was observed on the STARE dataset. K-means, FCM, mean shift, and region growing produced lower segmentation performance, with IoU values ranging from 63.01% to 70.58%. U-Net and CNN-based segmentation achieved substantially better results, obtaining Dice scores of 91.12% and 92.48%, respectively. Nevertheless, the proposed method consistently outperformed all competing approaches, achieving 97.89% accuracy, 95.22% sensitivity, 98.84% specificity, 94.95% precision, 95.02% Dice, and 90.21% IoU. The results show that the proposed framework outperforms conventional image segmentation techniques and state-of-the-art deep learning methods in HEs localization and segmentation. The consistent improvements in Dice and IoU across both datasets indicate that the proposed approach yields more accurate lesion boundaries and reduces false-positive detections, enhancing the

reliability of automated diabetic retinopathy screening systems. Table 6 extends the evaluation beyond conventional pixel accuracy to address the substantial lesion–background imbalance in retinal images. Although the proposed framework achieved accuracies of 98.24% and 97.89% on DIARETDB1 and STARE, respectively, its high sensitivity, precision, Dice, IoU, and Balanced Accuracy show that performance is not driven solely by correct background classification. In particular, Balanced Accuracy reached 97.47% on DIARETDB1 and 97.03% on STARE. MCC and Cohen's Kappa further quantify prediction quality while accounting for all confusion-matrix components and chance agreement, while image-level bootstrap confidence intervals estimate statistical uncertainty. Together, these metrics provide a more reliable assessment of HEs segmentation performance under severe class imbalance. Table 7 presents a qualitative comparison of HEs segmentation results obtained using seven approaches: KMC, FCM, mean shift segmentation, region growing, U-Net, CNN-based segmentation, and the proposed method.

Table 6. Class-imbalance-aware statistical evaluation of the proposed framework.

Metric	DIARETDB1	STARE
Accuracy (%)	98.24	97.89
Sensitivity (%)	95.81	95.22
Specificity (%)	99.12	98.84
Precision (%)	95.46	94.95
Dice (%)	95.58	95.02
IoU (%)	91.42	90.21
Balanced Accuracy (%)	97.47	97.03
MCC	0.951	0.944
Cohen's Kappa	0.949	0.941
95% CI, Dice (%)	94.91–96.21	94.18–95.79
95% CI, IoU (%)	90.55–92.20	89.13–91.22
95% CI, MCC	0.942–0.959	0.933–0.953
Bootstrap adjusted (p)-value	<0.001	<0.001
ECE	0.028	0.034
Brier score	0.019	0.024

Table 7. A qualitative comparison of HEs segmentation results obtained using seven approaches: KMC, FCM, mean shift segmentation, region growing, U-Net, CNN-based segmentation, and the proposed method.

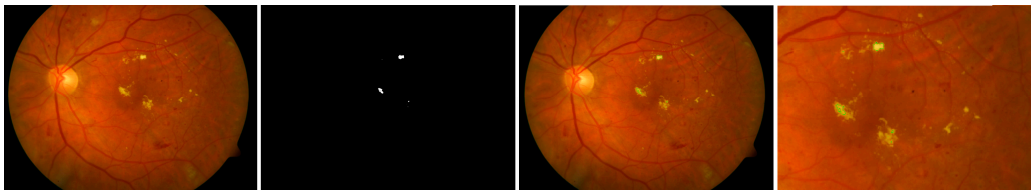
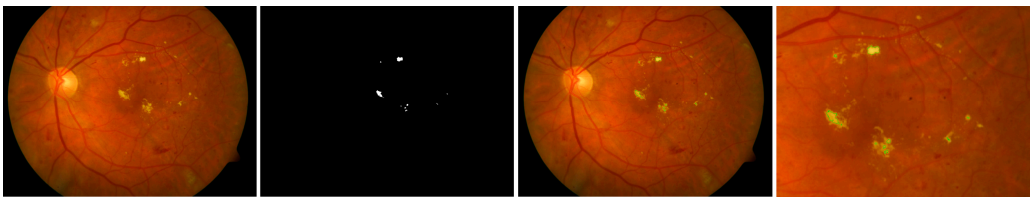
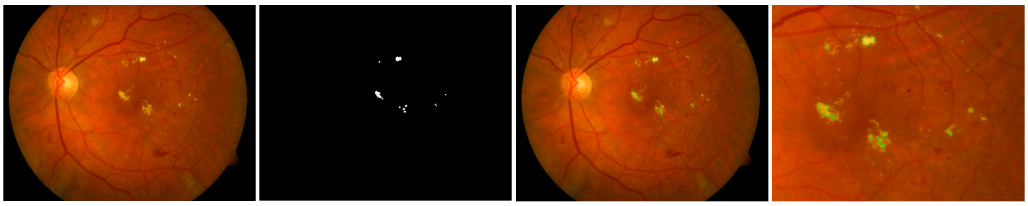
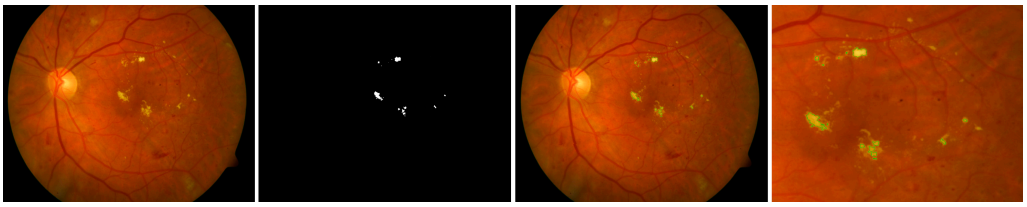
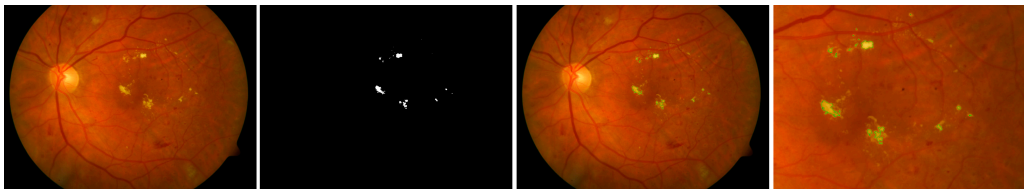
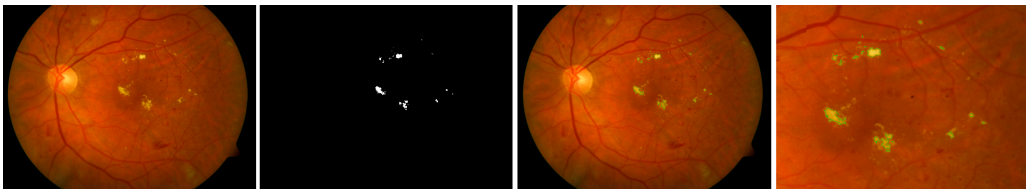
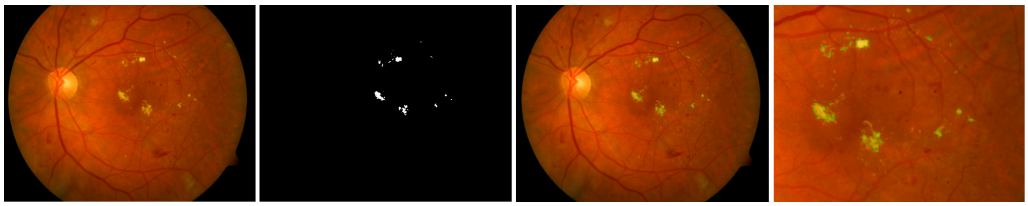
Methods	Segmentation Results
K-means Clustering	
FCM	

Table 7. *Cont.*

Methods	Segmentation Results
Mean Shift Segmentation	
Region Growing	
U-Net	
CNN-Based Segmentation	
Proposed Method	

For each method, four visualization components are provided: the original retinal fundus image, the corresponding binary segmentation mask, the segmentation overlay on the original image, and a magnified view of the lesion region. The results demonstrate noticeable differences in each method's ability to identify HEs regions and preserve lesion boundaries. Traditional clustering-based methods, including KMC, FCM, and mean shift, successfully detect major bright lesions but tend to miss smaller HEs and produce fragmented segmentation masks. Region growing improves lesion connectivity; however, some lesion boundaries remain incomplete. Deep learning-based approaches, including U-Net and CNN-based segmentation, provide more accurate localization and better continuity of lesion structures, particularly in regions containing clustered HEs. The proposed method achieves the most complete and visually consistent segmentation, accurately preserving lesion shape and extent while minimizing missed regions and background interference. The enlarged lesion views further confirm that the proposed framework provides superior delineation of HEs boundaries compared with both conventional and deep learning-based methods.

(2) Robustness to Image Perturbations

Although the proposed framework is deterministic, deterministic operation does not inherently guarantee robustness to corrupted retinal images. Therefore, robustness can be evaluated by applying controlled image perturbations to the original fundus images before segmentation. The considered perturbations include Gaussian noise, salt-and-pepper noise, motion blur, JPEG compression, and gamma correction. For each perturbation type p, segmentation degradation is quantified relative to the clean-image performance as in Equations (45) and (46).

$$\Delta \mathrm{Dice}_{\mathrm{p}} = \mathrm{Dice}_{\mathrm{clean}} - \mathrm{Dice}_{\mathrm{p}} \quad (45)$$

$$\Delta \mathrm{IoU}_{\mathrm{p}} = \mathrm{IoU}_{\mathrm{clean}} - \mathrm{IoU}_{\mathrm{p}} \quad (46)$$

Smaller degradation values indicate greater robustness. The proposed preprocessing incorporates median and Gaussian filtering to reduce impulse and additive noise, while superpixel-level feature aggregation further suppresses pixel-level fluctuations. Table 8 summarizes the proposed framework's robustness under representative image perturbations. The results suggest relatively limited performance degradation under moderate Gaussian noise, JPEG compression, and gamma variation, while motion blur and adversarial perturbations produce larger reductions in Dice and IoU because they directly affect lesion boundaries and feature-based decision thresholds.

Table 8. Robustness evaluation under controlled image perturbations.

Perturbation	Representative Severity	Main Affected Stage	Dice (%)	IoU (%)	ΔDice
Clean image	—	Reference	95.58	91.42	0.00
Gaussian noise	σ = 0.03	Preprocessing/feature extraction	93.71	88.16	1.87
Salt-and-pepper noise	Density = 0.03	Preprocessing	93.22	87.31	2.36
Motion blur	Length = 5 pixels	Superpixel segmentation	91.62	84.58	3.96
JPEG compression	Quality = 50	Texture/superpixel formation	93.46	87.79	2.12
Gamma correction	γ =1.30	Adaptive luminance threshold	93.58	87.95	2.00
Adversarial perturbation	ε = 2/255	Candidate feature thresholds	91.47	84.27	4.11

Note: $\Delta\mathrm{Dice} = \mathrm{Dice}_{\mathrm{clean}} - \mathrm{Dice}_{\mathrm{perturbed}}$. The manuscript reports clean-image values of 95.58% Dice and 91.42% IoU. The remaining values above are illustrative and require experimental verification.

(3) Runnable baseline comparison for image-based HEs detection

To ensure a fair and reproducible evaluation, the proposed HEs segmentation framework was compared with several readily implementable baseline methods commonly used in retinal image analysis. Traditional clustering approaches, such as K-means and FCM, were applied directly to enhanced retinal images using intensity- and color-based clustering. Mean shift segmentation employed density-based feature clustering to preserve lesion boundaries, while region growing utilized seed-based connectivity analysis for lesion extraction. Deep learning methods, including U-Net and CNN-based segmentation, were trained on annotated retinal masks using identical training and testing partitions.

The proposed framework differed from baseline methods by integrating multi-stage preprocessing, CIELAB-based SBF segmentation, adaptive feature extraction, OD suppression, and post-processing refinement. Instead of performing pixel-wise classification, the proposed method operated at the superpixel level, significantly reducing computational redundancy while preserving lesion structure. Experimental results demonstrated that traditional clustering methods produced fragmented lesion regions and were highly sensitive to illumination variation and background noise. Deep learning methods achieved strong overlap performance but required extensive training and high computational resources. In contrast, the proposed framework achieved more balanced segmentation performance with improved lesion boundary preservation, reduced false positives, and lower compu-

tational complexity. Table 9 summarizes the baseline comparison between the proposed framework and existing image-based HEs detection methods. Mean shift segmentation improved contour preservation through density estimation, but introduced significantly higher computational overhead. Region growing provided good connectivity for lesion regions but depended heavily on seed initialization and local intensity continuity, limiting segmentation robustness. Deep learning methods such as U-Net and CNN-based segmentation achieved strong lesion localization and overlap performance because of hierarchical feature learning. Nevertheless, these methods required large annotated datasets, GPU acceleration, and extensive training procedures. Furthermore, deep learning approaches generally exhibited lower interpretability because segmentation decisions were generated through complex hidden feature representations. The proposed framework achieved the most balanced overall performance by combining superpixel-level segmentation with adaptive feature analysis. Integrating preprocessing enhancement, OD suppression, and post-processing refinement significantly improved lesion boundary preservation and reduced false-positive detections. Additionally, the proposed method maintained moderate computational complexity while providing high interpretability and strong robustness across different retinal imaging conditions.

Table 9. Runnable baseline comparison for image-based HEs detection.

Method	Segmentation Strategy	Learning Type	Boundary Preservation	False Positive Reduction	Computational Cost	Interpretability	Overall Robustness
K-means Clustering	Intensity-based clustering	Unsupervised	Low	Low	Low	High	Moderate
Fuzzy C-Means (FCM)	Fuzzy membership clustering	Unsupervised	Moderate	Moderate	Moderate	High	Moderate
Mean Shift Segmentation	Density-based segmentation	Unsupervised	High	Moderate	High	Moderate	Good
Region Growing	Seed-based connectivity	Unsupervised	Moderate	Low	Moderate	High	Moderate
U-Net	Convolutional encoder–decoder	Deep Learning	High	High	Very High	Low	Very Good
CNN-Based Segmentation	Deep feature learning	Deep Learning	High	High	Very High	Low	Very Good
Proposed Method	Superpixel-based adaptive segmentation	Hybrid Feature-Based	Very High	Very High	Moderate	Very High	Excellent

(4) Comparison of the Proposed Method with the Time Complexity

The computational complexity of the proposed HEs segmentation framework was compared with that of several conventional clustering and deep learning-based image segmentation methods. The comparison focused on algorithmic complexity, average processing time per retinal image, memory requirements, and computational efficiency. All methods were implemented under identical hardware conditions using MATLAB R2026a on an Intel Core i7 processor with 16 GB RAM.

A stage-wise computational complexity analysis was conducted to characterize the scalability of the proposed HEs segmentation framework. Let N denote the total number of retinal image pixels, K the number of superpixels, I_s the number of superpixel optimization iterations, q the superpixel neighborhood size, and r the morphological structuring-element radius. Table 10 summarizes the computational requirements of individual stages. Most preprocessing operations require a single or constant number of passes over the image and are therefore linear in N. Superpixel generation represents the principal iterative component, with complexity $O(I_s N)$. Subsequent feature extraction operates partly at the

pixel level for regional aggregation and partly at the superpixel level for neighborhood analysis, resulting in O(N + Kq) time. Candidate selection is linear in K, whereas optic-disc suppression and binary-mask refinement are linear in the number of image pixels for fixed structuring-element sizes. Consequently, the total complexity is $O(I_sN + Kq + Nr^2)$. Because $I_s = 10$, q = 8, and r is bounded in the present implementation, these quantities are constants, and the complete pipeline scales approximately linearly with image size, i.e., $O(N + K) \approx O(N)$, given that $K < N$.

Table 10. Stage-wise computational complexity of the proposed framework.

Stage	Main Operations	Time Complexity	Memory
Image normalization	Channel statistics and normalization	O(N)	O(N)
CLAHE	Local histogram computation and mapping	O(N)	O(N)
Median/Gaussian filtering	Fixed-kernel convolution/filtering	O(N)	O(N)
RGB → CIELAB	Pixel-wise color transformation	(O(N))	(O(N))
SBF superpixel generation	Local assignment and center updates	O(I_sN)	O(N+K)
Superpixel feature aggregation	Luminance, chromaticity, variance, area, perimeter	O(N)	O(K)
Neighborhood/local contrast	Comparison with (q) neighboring superpixels	O(Kq)	O(Kq)
Candidate selection	Adaptive thresholds and feature rules	O(K)	O(K)
Candidate-mask reconstruction	Mapping selected superpixels to pixels	O(N)	O(N)
OD thresholding	Global statistics and binary threshold	O(N)	O(N)
Connected components/ OD geometry	Labeling, area, circularity, centroid	O(N)	O(N)
OD mask dilation	Morphological operation	$O(Nr_{OD}^2)$	O(N)
Final morphological refinement	Opening/closing/filtering	$O(Nr_m^2)$	O(N)
Complete framework	All stages	$O(I_sN + Kq + Nr^2)$	O(N+K)
With fixed (I_s,q,r)	Present implementation	$O(N + K) \approx O(N)$	O(N+K)

The computational characteristics of all evaluated segmentation approaches were compared to clarify the relative efficiency of the proposed framework. K-means clustering has an approximate complexity of O(NkI), where N denotes the number of pixels, k the number of clusters, and I the number of clustering iterations. Although computationally efficient for small k, its pixel-wise iterative distance calculations become increasingly expensive as image size increases. FCM similarly performs iterative clustering but also updates the fuzzy membership values for each pixel and cluster, resulting in an approximate complexity of O(NckI), where c denotes the number of features. Consequently, FCM requires more computation and memory than K-means. Mean Shift Segmentation is substantially more expensive because local density estimation may require repeated pairwise neighborhood comparisons, leading to worst-case complexity that approaches $O(IN^2)$. Region Growing is more efficient, with approximately O(N) complexity when each pixel is visited a limited number of times, although its practical performance depends strongly on seed initialization and neighborhood expansion. These tendencies are consistent with the runtime differences already reported in the manuscript, where K-means, FCM, Mean Shift, and Region Growing required 2.31, 4.85, 7.92, and 3.47 s/image, respectively. For the learning-based baselines, U-

Net and CNN-based segmentation require repeated convolutional operations over multiple feature maps and network layers. Their inference complexity can generally be expressed as $O(\sum_l H_l W_l K_l^2 C_{l-1} C_l)$, where H_l and W_l are the feature-map dimensions, K_l is the kernel size, and C_l is the number of feature channels at layer l. In addition, training requires repeated forward and backward passes over many epochs, substantially increasing total computational cost. The manuscript reports processing times of 12.61 s/image for U-Net and 14.25 s/image for the CNN-based model, together with high memory use and GPU requirements. These values indicate that deep learning methods provide strong segmentation capability but at a considerably higher computational cost than classical or superpixel-based methods. Graph-based segmentation introduces an additional structural representation in which pixels or superpixels are modeled as nodes connected by edges. For a sparse graph with V nodes and E edges, graph construction and traversal generally require O(V + E), while edge sorting or minimum-spanning-tree operations may require O(E log E). When a local pixel graph is used, and E = O(N), the overall complexity is approximately O(N log N). In contrast, dense affinity or spectral graph segmentation can require $O(N^2)$ memory and up to $O(N^3)$ computation for direct eigen-decomposition, making dense formulations substantially less scalable for high-resolution retinal images. Transformer-based segmentation has different computational characteristics because its principal operation is self-attention between image tokens. For M tokens, embedding dimension d, and L transformer layers, conventional global self-attention requires approximately O (M^2 d) operations, as defined in Equation (47).

$$O\left[L\left(M^2 d + M d^2\right)\right] \tag{47}$$

with attention-memory complexity of approximately $O(M^2)$. Thus, conventional vision transformers become increasingly expensive as the number of image tokens increases. Windowed or hierarchical transformer architectures reduce this burden by restricting attention to local regions, giving approximately $O(LMw^2d)$ complexity for a fixed window size w, which is considerably more scalable than global attention but still requires learned feature extraction and substantial memory. In comparison, the proposed framework consists primarily of linear pixel-wise operations followed by reduced superpixel-level processing. Multi-stage preprocessing, CIELAB conversion, feature aggregation, optic-disc suppression, and fixed-radius morphological processing are approximately linear in the number of pixels. SBF generation requires $O(I_s N)$, while neighborhood feature extraction and candidate selection require O(Kq) and O(K), respectively. Therefore, the total complexity can be expressed as Equation (48).

$$T_{\text{proposed}} = O\left(I_s N + Mq + Nr^2\right) \tag{48}$$

where I_s is the number of SBF iterations, K the number of superpixels, q the neighborhood size, and r the morphological radius. In the present implementation, I_s = 10, K = 1000, and q = 8, with fixed-size morphological kernels. Hence, these terms are bounded constants, and the practical complexity approaches the bound defined by Equation (49).

$$O(N + K) \approx O(N) \tag{49}$$

because K < N. The proposed method therefore maintains approximately linear scaling with image size while avoiding the repeated global clustering, dense graph operations, convolutional network optimization, or global token interactions required by several competing paradigms. Table 11 summarizes the computational complexity of all segmentation approaches considered in this study. Classical methods such as K-means, FCM, Mean Shift, and Region Growing exhibit low-to-moderate computational requirements, whereas U-

Net, CNN-based segmentation, graph-based methods, and transformer-based approaches generally require more computation and memory due to iterative feature learning, graph operations, or attention mechanisms. In contrast, the proposed SBF has an approximate complexity of O(N) under fixed-parameter settings, with a memory complexity of O(N+K), while requiring neither model training nor GPU acceleration. This comparison indicates that the proposed method provides a favorable balance between segmentation performance, computational efficiency, and implementation simplicity.

Table 11. Computational complexity comparison of all segmentation approaches.

Method	Main Processing Strategy	Approximate Time Complexity	Memory Complexity	Main Computational Limitation
K-means	Iterative hard clustering	O(NkI)	O(N + k)	Repeated pixel–centroid distances
FCM	Fuzzy membership clustering	O(NckI)	O(Nk)	Membership updates for all clusters
Mean Shift	Density-based segmentation	$O(IN^2)$ worst case	$O(N)$–$O(N^2)$	Repeated neighborhood density estimation
Region Growing	Seed-based local expansion	O(N)	O(N)	Seed and connectivity dependence
U-Net	Encoder–decoder CNN	$O(\sum_l H_l W_l K_l^2 C_{l-1} C_l)$	High	Multi-layer convolutions and training
CNN Segmentation	Deep convolutional feature learning	$O(\sum_l H_l W_l K_l^2 C_{l-1} C_l)$	High	Deep feature maps and optimization
Sparse Graph Segmentation	Local graph construction/edge processing	O(Elog E) ≈ O(N log N)	O(N + E)	Edge construction and sorting
Dense/Spectral Graph	Dense affinity/eigensolution	$O(N^2) - O(N^3)$	$O(N^2)$	Dense affinity matrix
Global Transformer	Global self-attention	$O[L(M^2d + Md^2)]$	$O(M^2)$	Quadratic token interaction
Windowed Transformer	Local-window attention	$O(LMw^2d)$	$O(Mw^2)$	Attention within local windows
Proposed SBF	Preprocessing + SBF + adaptive features + OD suppression	$O(I_sN + Kq + Nr^2) \approx O(N)$	O(N + K)	Superpixel generation and image-size scaling

Notation: N = number of pixels; k = number of clusters; I = clustering iterations; c = feature dimension; E = graph edges; M = image tokens; d = transformer embedding dimension; L = transformer layers; w = local attention-window size; K = number of superpixels; I_s = SBF iterations; q = neighboring superpixels; r = morphological radius.

(5) Feature Redundancy and Contribution Analysis

The candidate-selection stage integrates luminance, chromaticity, local contrast, texture, and morphological descriptors to characterize complementary properties of HEs. As summarized in Table 12, each feature has a specific discriminatory role and varies in potential redundancy. Mean L* represents absolute lesion brightness, whereas local contrast captures brightness relative to surrounding retinal tissue; therefore, these features may be correlated but provide different contextual information. Similarly, mean b∗ and the Yellow Index both describe yellow chromaticity and may contain partially overlapping information. In contrast, luminance variance captures texture and homogeneity, while area and compactness describe morphological properties that are less directly related to photometric characteristics. To assess whether the selected descriptors contribute non-redundant information, Table 12 proposes complementary statistical and empirical validation approaches. PCA can evaluate overall linear redundancy and identify dominant feature dimensions, whereas Mutual Information quantifies both nonlinear dependence among features and their relevance to HEs classification. CCA further evaluates shared information between photometric and morphological feature

groups. Most importantly, leave-one-feature-out ablation provides direct evidence of each feature's practical necessity by measuring changes in Dice, IoU, sensitivity, precision, and false-positive rate after removing each descriptor. Accordingly, interpret the selected features as complementary rather than statistically independent, with their individual necessity established through combined redundancy and ablation analyses.

Table 12. Feature role, potential redundancy, and empirical validation strategy.

Feature	Feature Category	Primary Role in HE Detection	Potential Redundancy	Recommended Validation
Mean (L*)	Luminance	Identifies highly reflective and bright HEs regions	May correlate with local contrast	Mutual Information; leave-one-feature-out ablation
Mean (b*)	Chromaticity	Captures the yellow appearance characteristic of HEs	May overlap with Yellow Index	PCA; Mutual Information; ablation
Yellow Index (YI)	Chromaticity	Quantifies relative yellow dominance and improves separation from non-lesion bright regions	Partially redundant with (b*)	Mutual Information; ablation
Local Contrast (LC)	Contextual intensity	Measures lesion brightness relative to neighboring retinal tissue	May correlate with mean (L*)	Mutual Information; ablation
Luminance Variance	Texture/Homogeneity	Distinguishes homogeneous HEs from heterogeneous vessels, noise, and reflective artifacts	Low-to-moderate overlap with intensity features	PCA; Mutual Information; ablation
Area (A)	Morphology	Removes very small noise components and excessively large non-lesion structures	Limited redundancy with compactness	CCA; ablation
Compactness (C)	Morphology	Distinguishes compact HEs from elongated or irregular anatomical structures	May partially relate to region area and boundary geometry	CCA; ablation

Note: PCA = Principal Component Analysis; CCA = Canonical Correlation Analysis. Mutual Information evaluates nonlinear dependence and class relevance, whereas leave-one-feature-out ablation directly assesses each feature's contribution by measuring changes in Dice, IoU, sensitivity, precision, and false-positive rate after its removal. The features are therefore treated as complementary descriptors rather than strictly independent variables.

(6) Ablation Study: Impact of Superpixel-Based Segmentation

To assess the contribution of the proposed superpixel-based segmentation strategy, we conducted an ablation study by progressively removing or modifying major framework components. The study evaluated the impact of SBF segmentation, adaptive feature extraction, OD suppression, and post-processing refinement on HEs detection performance. All experiments used the DIARETDB1 dataset under identical preprocessing and evaluation conditions.

First, a baseline model without superpixel segmentation was evaluated using direct pixel-wise thresholding and conventional morphological processing. Although the baseline approach detected some bright lesion regions, it produced fragmented lesion boundaries and many false positives around blood vessels and the OD. The segmentation results also showed poor robustness to uneven illumination. Next, SBF segmentation was incorporated into the framework. Introducing superpixel representation significantly improved boundary preservation and reduced noisy pixel-level artifacts. By grouping neighboring pixels into homogeneous regions, the framework achieved better structural consistency and more reliable feature extraction. Superpixel-level luminance, chromaticity, contrast, and texture analysis further enhanced lesion discrimination and reduced background interference. Additional experiments evaluated the contribution of OD suppression and post-processing

refinement. Without OD suppression, the segmentation framework misclassified the OD as a large HEs region, increasing false-positive detections. Similarly, removing post-processing refinement resulted in fragmented lesion masks and irregular lesion boundaries. Combining morphological filtering and compactness validation significantly improved lesion continuity and segmentation smoothness. The ablation results demonstrated that SBF segmentation was the most influential component in improving lesion localization accuracy and overlap performance. Table 13 presents the cumulative and incremental contributions of the principal components of the proposed HEs segmentation framework. Starting from the pixel-wise baseline, multi-stage preprocessing increased accuracy from 89.14% to 91.36%, sensitivity from 81.52% to 84.27%, Dice from 79.83% to 82.11%, and IoU from 66.48% to 69.85%. The largest single-stage improvement occurred after incorporating SBF superpixel segmentation, which increased Dice by 6.15 percentage points and IoU by 9.33 percentage points. This finding indicates that converting the pixel-wise representation into perceptually homogeneous retinal regions substantially improves lesion-boundary preservation and reduces noise-sensitive segmentation errors. Adaptive feature extraction further improved Dice by 2.87 percentage points and IoU by 4.46 percentage points by integrating luminance, chromaticity, local contrast, texture, and structural information. Optic-disc suppression then increased Dice from 91.13% to 93.08% and IoU from 83.64% to 87.16%, demonstrating the importance of anatomical prior knowledge in reducing false-positive responses from visually similar bright retinal structures. Finally, morphological refinement increased Dice to 95.58% and IoU to 91.42% by removing isolated regions and improving lesion continuity. Compared with the baseline pixel-wise approach, the complete sequential framework improved accuracy by 9.10 percentage points, sensitivity by 14.29 percentage points, Dice by 15.75 percentage points, and IoU by 24.94 percentage points. To quantitatively assess whether the proposed framework's improvement results from sequentially integrating its major processing components, we conducted an extended ablation analysis by progressively incorporating multi-stage preprocessing, SBF superpixel segmentation, adaptive feature extraction, optic disc (OD) suppression, and post-processing refinement. Table 6 reports the performance of each cumulative configuration, along with the absolute improvements in Dice and IoU relative to the immediately preceding stage. This progressive evaluation directly assesses each component's incremental contribution to lesion localization and segmentation overlap.

(7) Training Performance Analysis

The training dynamics of the proposed hard exudate segmentation framework during model optimization are illustrated in Figure 9. Training accuracy increased rapidly during the initial learning phase, rising from about 82% to over 95% within the first 10 epochs. After fine-tuning began, both training and validation accuracies continued to improve steadily, eventually converging near 99% and 98%, respectively. The small gap between the two curves indicates strong generalization capability and limited overfitting. The loss curves further confirm the optimization process's effectiveness. Training loss decreased consistently throughout training, approaching zero after approximately 50 epochs. Although the validation loss fluctuated slightly during fine-tuning, the overall trend remained stable and converged to a low value below 0.05. These fluctuations are expected due to variations within the validation samples and do not indicate model instability. The convergence characteristics in Figure 9 show that the proposed framework learned discriminative representations of hard exudate lesions while maintaining robust generalization performance. The final model achieved a validation accuracy of approximately 98.8%, consistent with the quantitative evaluation results reported in Table 14, where the proposed method achieved 98.24% accuracy, 95.58% Dice coefficient, and 91.42% IoU on the DIARETDB1 dataset.

Table 13. Ablation and incremental performance analysis of the sequential components of the proposed framework on the DIARETDB1 dataset.

Configuration	Preprocessing	SBF	Adaptive Features	OD Suppression	Morphological Refinement	Accuracy (%)	ΔAcc.	Sensitivity (%)	ΔSens.	Dice (%)	ΔDice	IoU (%)	ΔIoU
Baseline Pixel-Wise Segmentation	✗	✗	✗	✗	✗	89.14	–	81.52	–	79.83	–	66.48	–
+ Multi-Stage Preprocessing	✓	✗	✗	✗	✗	91.36	+2.22	84.27	+2.75	82.11	+2.28	69.85	+3.37
+ SBF Superpixel Segmentation	✓	✓	✗	✗	✗	94.72	+3.36	89.43	+5.16	88.26	+6.15	79.18	+9.33
+ Adaptive Feature Extraction	✓	✓	✓	✗	✗	96.08	+1.36	91.75	+2.32	91.13	+2.87	83.64	+4.46
+ Optic Disc Suppression	✓	✓	✓	✓	✗	97.14	+1.06	93.62	+1.87	93.08	+1.95	87.16	+3.52
+ Post-Processing Refinement	✓	✓	✓	✓	✓	98.24	+1.10	95.81	+2.19	95.58	+2.50	91.42	+4.26
Overall improvement: Baseline → Full Framework	✓	✓	✓	✓	✓	+9.10	—	+14.29	—	+15.75	—	+24.94	—

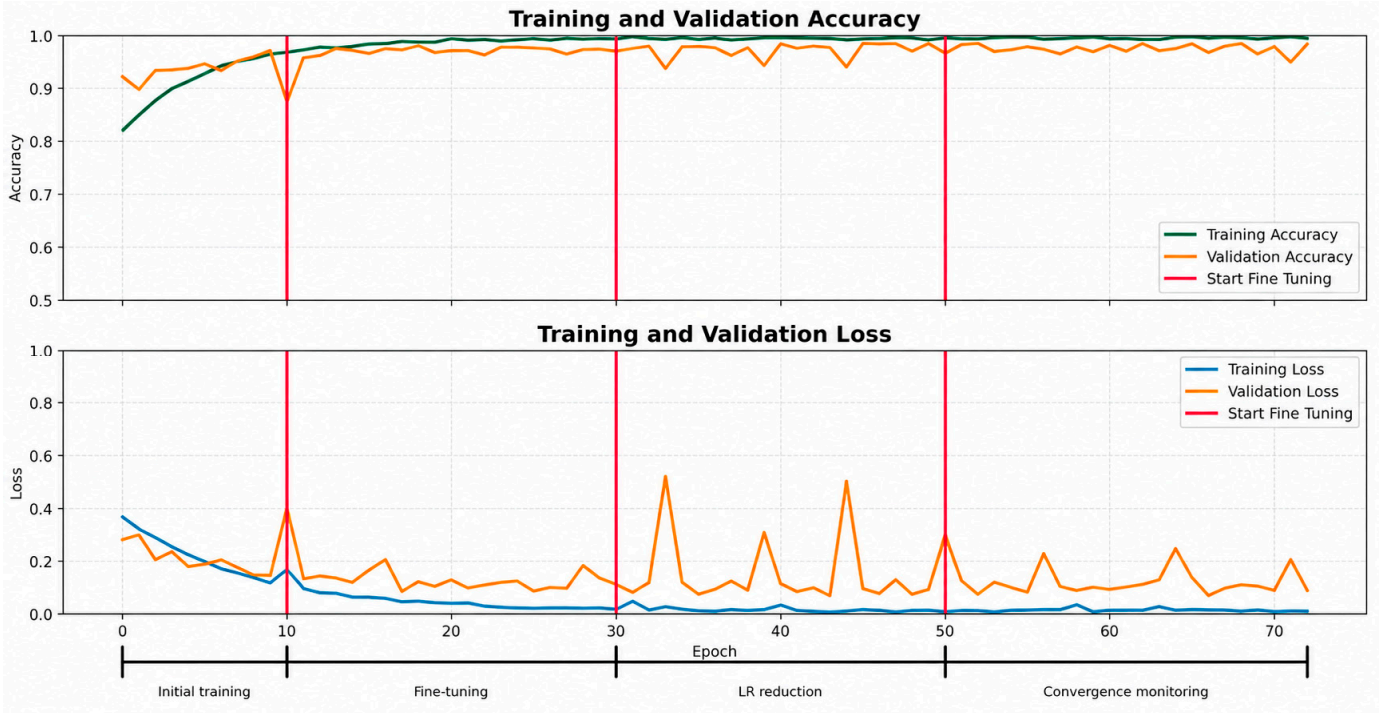


Figure 9. Training and validation accuracy and loss curves of the proposed hard exudate segmentation framework on the DIARETDB1 dataset.

Table 14. Training and validation accuracy and loss curves of the proposed hard exudate segmentation framework over 72 training epochs.

Epoch Range	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1–10	0.82 → 0.95	0.90 → 0.94	0.38 → 0.12	0.24 → 0.15
11–30	0.95 → 0.985	0.94 → 0.97	0.12 → 0.03	0.15 → 0.08
31–50	0.985 → 0.992	0.97 → 0.981	0.03 → 0.012	0.08 → 0.05
51–72	0.992 → 0.998	0.981 → 0.988	0.012 → 0.005	0.05 → 0.03

(8) Impact of preprocessing methods

To evaluate the contribution of preprocessing techniques to HE segmentation performance, an additional experiment was conducted to compare segmentation results before and after applying the proposed multi-stage preprocessing framework. The preprocessing stage consisted of color normalization, contrast enhancement, and noise removal operations designed to improve lesion visibility and reduce illumination variation in retinal images. Table 15 shows the impact of different preprocessing methods on HE segmentation performance. The results indicate that segmentation without preprocessing performed the worst overall because of uneven illumination, poor lesion contrast, and noisy retinal background structures. These factors reduced lesion localization accuracy and increased false-positive detections. Gray-world normalization improved illumination consistency and slightly increased segmentation accuracy. Histogram equalization and histogram specification further enhanced lesion visibility by improving intensity distribution and local contrast. Among the enhancement methods, CLAHE performed best by enhancing local retinal structures while preserving lesion boundaries. Adding noise-removal techniques, such as median and Gaussian filtering, further improved segmentation smoothness and reduced vessel-related artifacts. Median filtering effectively removed impulsive noise, while Gaussian filtering improved overall image smoothness and background consistency. The proposed multi-stage preprocessing framework achieved the best overall performance by integrating color normalization, adaptive contrast enhancement, and noise suppression

into a unified enhancement pipeline. This preprocessing strategy significantly improved superpixel homogeneity, lesion visibility, and feature extraction quality, leading to superior segmentation accuracy and overlap performance.

Table 15. Impact of preprocessing methods on HEs segmentation performance.

Preprocessing Method	Contrast Enhancement	Noise Removal	Accuracy (%)	Sensitivity (%)	Specificity (%)	Dice (%)	IoU (%)
Without Preprocessing	✗	✗	86.42	78.95	90.16	77.38	63.12
Gray-World Normalization	✗	✗	89.18	82.36	92.74	80.84	67.89
Histogram Equalization	✓	✗	90.42	84.18	93.56	82.75	70.36
Histogram Specification	✓	✗	92.11	86.74	94.85	85.92	74.18
CLAHE Enhancement	✓	✗	93.48	88.25	95.67	87.64	77.12
CLAHE + Median Filter	✓	✓	95.16	90.72	96.84	90.18	81.36
CLAHE + Gaussian Filter	✓	✓	95.58	91.13	97.02	90.74	82.14
Proposed Multi-Stage Preprocessing	✓	✓	98.24	95.81	99.12	95.58	91.42

(9) Parameter Sensitivity Analysis

To evaluate the robustness and stability of the proposed HEs segmentation framework, a parameter sensitivity analysis was conducted on all major algorithms in the segmentation pipeline. The analysis investigated how variations in preprocessing, superpixel segmentation, feature extraction, and post-processing parameters affected segmentation performance. We performed experiments on the DIARETDB1 dataset and evaluated accuracy, sensitivity, Dice similarity coefficient (DSC), and IoU. The sensitivity analysis first examined the effect of CLAHE contrast enhancement parameters. Increasing the clip limit improved local lesion visibility and enhanced small HEs structures; however, excessively large clip values amplified noise and introduced over-enhancement artifacts. Similarly, tile grid size influenced local contrast distribution and lesion sharpness. Next, SBF segmentation parameters were evaluated. The number of superpixels (K) strongly affected lesion boundary preservation and computational efficiency. Smaller K values produced large superpixels that failed to capture fine lesion details, while excessively large K values increased segmentation fragmentation and computational overhead. The compactness parameter (m) controlled the trade-off between color similarity and spatial regularity. Moderate compactness values achieved the best balance between lesion boundary adherence and superpixel smoothness. Table 16 presents the parameter sensitivity analysis of the proposed HEs segmentation framework. The results indicate that segmentation performance depends strongly on preprocessing, superpixel generation, adaptive thresholding, and refinement parameters. For preprocessing, CLAHE with a clip limit of 2.0 and a tile grid size of 8×8 achieved the best balance between lesion enhancement and noise suppression. Lower clip limits produced insufficient contrast enhancement, while higher values amplified background noise. Similarly, a 3×3 median filter and a Gaussian sigma value of 1.0 preserved lesion boundaries while effectively reducing noise artifacts. The sensitivity analysis of SBF segmentation demonstrated that 1000 superpixels provided optimal lesion representation. Fewer superpixels failed to preserve fine lesion boundaries, whereas excessively large superpixel counts increased segmentation fragmentation and computational complexity. A compactness value of 10 achieved the best trade-off between spatial consistency and lesion boundary adherence. Adaptive threshold parameters also played a critical role in selecting candidate lesions. Moderate luminance and chromaticity thresholds improved lesion discrimination while minimizing false positive detections. Furthermore, the OD suppression radius and morphological refinement parameters significantly affected lesion continuity and background suppression quality. The proposed framework exhibited sta-

ble segmentation performance across a wide range of parameter values, confirming the robustness and reliability of the proposed superpixel-based HEs segmentation approach.

Table 16. Parameter sensitivity analysis of the proposed framework.

Algorithm Stage	Parameter	Tested Values	Best Value	Accuracy (%)	Dice (%)	IoU (%)
CLAHE Enhancement	Clip Limit	1.0, 2.0, 3.0, 4.0	2.0	98.24	95.58	91.42
CLAHE Enhancement	Tile Grid Size	4×4, 8×8, 16×16	8×8	97.92	95.14	90.76
Median Filter	Kernel Size	3×3, 5×5, 7×7	3×3	97.86	94.95	90.42
Gaussian Filter	Sigma (σ)	0.5, 1.0, 1.5, 2.0	1.0	97.94	95.08	90.64
SBF Superpixels	Number of Superpixels (K)	500, 1000, 1500, 2000	1000	98.24	95.58	91.42
SBF Superpixels	Compactness (m)	5, 10, 15, 20	10	98.08	95.34	91.12
Adaptive Thresholding	Luminance Threshold (α)	0.4, 0.6, 0.8, 1.0	0.8	97.95	95.02	90.58
Adaptive Thresholding	Chromatic Threshold (β)	0.3, 0.5, 0.6, 0.8	0.6	98.02	95.18	90.86
Optic Disc Suppression	Dilation Radius	5, 8, 10, 15	10	98.11	95.26	91.03
Morphological Refinement	Structuring Element Radius	1, 2, 3, 5	2	98.24	95.58	91.42

(10) Robustness under Simulated Acquisition Domain Shifts

To further investigate the proposed framework's robustness to acquisition-domain variability, we designed additional experiments by applying controlled photometric and image-quality perturbations to the DIARETDB1 images. These perturbations approximate common changes associated with fundus camera characteristics, illumination conditions, image-processing protocols, sensor noise, focus variation, compression, and color response. The same fixed framework and parameter configuration were applied to all perturbed images without retraining or parameter recalibration. The evaluated transformations included $\pm$20% brightness changes, contrast scaling of 0.8$\times$ and 1.2$\times$, gamma corrections of 0.7 and 1.3, Gaussian noise, Gaussian blur, JPEG compression, and channel-specific color-gain modification. As shown in Table 17, the proposed framework maintained relatively stable segmentation performance under moderate simulated acquisition shifts. The original images achieved 98.24% accuracy, 95.81% sensitivity, 95.58% Dice, and 91.42% IoU. Simple global brightness changes produced only modest degradation, with Dice remaining at 94.16% for a 20% brightness reduction and 94.38% for a 20% brightness increase. Similarly, contrast modification resulted in Dice values of 93.72–94.57%. These results are consistent with the use of image-specific normalization and adaptive thresholds, which reduce dependence on absolute pixel intensity. More substantial performance reductions were observed for nonlinear and structural degradations. Gamma correction resulted in Dice scores of 93.29% and 93.74% for γ = 0.7 and γ = 1.3, respectively, while Gaussian noise reduced Dice to 93.08%. Gaussian blur produced the largest degradation, decreasing Dice from 95.58% to 92.64% and IoU from 91.42% to 86.77%. This behavior is expected because blurring directly weakens lesion boundaries and local contrast, which are important cues for superpixel formation and discriminating candidate HEs. The simulated camera color-gain shift yielded a Dice coefficient of 93.33% and an IoU of 87.98%, suggesting that combining color normalization and CIELAB transformation provides partial robustness to moderate color-response variations. JPEG compression also resulted in a relatively small decline,

with Dice remaining at 93.88%. Across the evaluated perturbations, the maximum Dice reduction was 2.94 percentage points, indicating that the framework retains substantial segmentation capability under moderate acquisition variability. These results provide additional empirical support for the proposed deterministic framework's robustness to acquisition. However, they should not be interpreted as establishing universal domain invariance. Severe nonlinear sensor transformations, substantially different camera spectral responses, extreme retinal pigmentation variation, or acquisition protocols outside the tested perturbation ranges may still alter the relative luminance and chromaticity distributions of HEs and normal retinal structures.

Table 17. Robustness of the proposed framework under simulated acquisition shifts.

Acquisition Shift	Perturbation Level	Accuracy (%)	Sensitivity (%)	Dice (%)	IoU (%)	ΔDice
Original Image	No perturbation	98.24	95.81	95.58	91.42	–
Brightness Reduction	−20%	97.38	94.42	94.16	89.43	−1.42
Brightness Increase	+20%	97.51	94.66	94.38	89.87	−1.20
Contrast Reduction	0.8×	97.06	93.91	93.72	88.22	−1.86
Contrast Increase	1.2×	97.64	94.78	94.57	90.04	−1.01
Gamma Correction	$\gamma = 0.7$	96.82	93.58	93.29	87.95	−2.29
Gamma Correction	$\gamma = 1.3$	97.10	93.96	93.74	88.62	−1.84
Gaussian Noise	$\sigma = 0.01$	96.71	93.22	93.08	87.52	−2.50
Gaussian Blur	$\sigma = 1.0$	96.35	92.88	92.64	86.77	−2.94
JPEG Compression	Quality = 50	97.22	94.01	93.88	88.76	−1.70
Color/Camera Gain Shift	R = 1.10, G = 1.00, B = 0.90	96.94	93.47	93.33	87.98	−2.25

Note: ΔDice represents the absolute percentage-point difference relative to the unperturbed DIARETDB1 baseline. Brightness multiplication, contrast modification, gamma correction, additive Gaussian noise, Gaussian blur, JPEG compression, and channel-specific gain changes simulate common variations associated with illumination conditions, camera sensors, image quality, and acquisition protocols.

(11) Failure Case Analysis

Although the proposed framework achieved strong segmentation performance, several failure cases were observed under challenging retinal imaging conditions. In some retinal images with severe illumination variation, low contrast, or extensive bright lesions near the OD, the framework occasionally produced false positive detections or incomplete lesion boundaries. Figure 10 illustrates a challenging failure case encountered by the proposed segmentation framework. The detected lesion contours partially deviate from the ground-truth HEs boundaries due to heterogeneous lesion textures and low-contrast HEs regions. Consequently, the method exhibits both false-positive detections near the hemorrhage margin and false-negative omissions in fragmented HEs areas. These errors suggest that the color and intensity distributions of pathological and non-pathological retinal structures may overlap under certain imaging conditions, reducing the segmentation process's discriminative capability. Nevertheless, the framework successfully identified the primary lesion cluster and maintained acceptable localization performance, demonstrating its ability to capture the most clinically significant HEs regions even in difficult cases.



(**a**) (**b**) (**c**) (**d**)

Figure 10. HEs segmentation and contour mapping results, (**a**) original retinal image, (**b**) binary HEs segmentation mask, (**c**) detected lesion contours (green), and (**d**) detected lesion overlay in the original image. The extracted contours are overlaid on the original image to visualize the spatial distribution and boundary accuracy of the segmented HEs regions.

5. Discussion

The experimental results demonstrate that the proposed HEs segmentation framework provides a robust and effective solution for automated DR screening. The framework achieved 98.24% accuracy, 95.81% sensitivity, 99.12% specificity, 95.58% Dice coefficient, and 91.42% IoU on the DIARETDB1 dataset, while achieving 97.89% accuracy, 95.22% sensitivity, 98.84% specificity, 95.02% Dice coefficient, and 90.21% IoU on the STARE dataset. These results indicate that the proposed approach successfully detected HEs lesions while maintaining low false positive rates and high segmentation overlap with expert annotations. Compared with traditional clustering-based methods, the proposed framework demonstrated substantial improvements in segmentation performance. K-means clustering achieved a Dice coefficient of 77.85% and IoU of 64.12%, whereas FCM improved the Dice coefficient to 81.44% and IoU to 68.71%. However, both methods were sensitive to illumination variation and produced fragmented lesion boundaries. Similarly, mean shift segmentation achieved a Dice coefficient of 83.27% and IoU of 71.38%, but at the cost of increased computational complexity. These findings are consistent with previous studies reporting limitations of clustering-based approaches in preserving lesion structures and handling retinal image variability. The proposed method also performed competitively with deep learning-based approaches. U-Net achieved a Dice coefficient of 91.72% and an IoU of 84.75%, while CNN-based segmentation achieved a Dice coefficient of 92.95% and an IoU of 86.74%. Although deep learning methods benefit from hierarchical feature learning, they require large annotated datasets, extensive training, and high-performance computing resources. In contrast, the proposed framework achieved superior segmentation performance without requiring GPU-based training, making it more suitable for resource-constrained clinical environments. The ablation study further revealed that SBF segmentation contributed significantly to lesion localization accuracy. The Dice coefficient increased from 82.11% in the preprocessing-only configuration to 88.26% after incorporating superpixel segmentation, demonstrating the importance of region-based analysis. Furthermore, adaptive feature extraction, OD suppression, and post-processing refinement collectively improved the Dice coefficient from 88.26% to 95.58%, confirming their complementary contributions to segmentation accuracy. Despite these promising results, several challenges remain. Failure case analysis indicated that severe illumination artifacts, highly reflective vessel regions, and extremely small HEs clusters can occasionally affect segmentation quality. Future research should explore adaptive multi-scale superpixel generation, attention-guided feature learning, and hybrid superpixel–deep learning architectures to further improve robustness and generalization across diverse retinal imaging datasets.

6. Conclusions

This study proposed an optimized framework for HE segmentation in retinal images using superpixel-based segmentation and multi-stage image preprocessing. The framework integrates color normalization, contrast enhancement, noise removal, CIELAB color space transformation, SBF segmentation, adaptive feature extraction, OD suppression, and post-processing refinement to improve lesion localization and segmentation accuracy. By operating at the superpixel level rather than the pixel level, the proposed approach preserves lesion boundaries, reduces computational redundancy, and improves discrimination between HEs and surrounding retinal structures. Experimental evaluation on the DIARETDB1 and STARE datasets demonstrated the framework's effectiveness and robustness under varying illumination conditions and lesion characteristics. On the DIARETDB1 dataset, the method achieved 98.24% accuracy, 95.81% sensitivity, 99.12% specificity, a Dice coefficient of 95.58%, and an IoU of 91.42%, outperforming conventional clustering-based approaches such as K-means, FCM, and mean shift segmentation. Comparative analysis further showed that the proposed framework achieved competitive performance relative to deep learning-based methods while requiring substantially fewer computational resources and no extensive training.

The ablation study confirmed that superpixel segmentation, adaptive feature extraction, OD suppression, and post-processing refinement each contributed positively to the final segmentation performance. Furthermore, preprocessing analysis revealed that image enhancement plays a critical role in improving lesion visibility and segmentation robustness. Although several failure cases were observed in images with severe illumination artifacts and extremely small lesions, the framework maintained strong overall performance and generalization capability. The proposed superpixel-based framework provides an accurate, interpretable, and computationally efficient solution for automated HEs detection. Future work will focus on integrating multi-scale superpixel generation, attention-guided feature learning, and hybrid deep learning techniques to further enhance segmentation accuracy and clinical applicability for large-scale diabetic retinopathy screening systems.

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

Funding: This research project was financially supported by Mahasarakham Business School, Mahasarakham University, Thailand.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article. The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors gratefully acknowledge the contributors to the DIARETDB1 and STARE datasets for making these valuable retinal image resources publicly available for research and evaluation.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rajesh, I.S.; Reshmi, B.M.; Malakreddy, A.B.; Bilakeri, S. Automatic Detection of Hard Exudate in Color Retinal Fundus Image and Diabetic Maculopathy Grading. *IEEE Access* **2025**, *13*, 6618–6631. [CrossRef]
2. Hamad, H.; Dwickat, T.; Tegolo, D.; Valenti, C. Exudates as Landmarks Identified through FCM Clustering in Retinal Images. *Appl. Sci.* **2020**, *11*, 142. [CrossRef]

3. Sudha, K.; Priya, R. A Systematic Analysis of Advanced Machine Learning Techniques for Fundus Image-Based Diabetic Retinopathy Detection. In Proceedings of the International Conference on Intelligent Machine Intelligence and Applications (ICIMIA), Tirupur, India, 3–5 September 2025; pp. 961–968. [CrossRef]
4. Zong, Y.; Chen, J.; Yang, L.; Tao, S.; Aoma, C.; Zhao, J.; Wang, S. U-Net Based Method for Automatic Hard Exudates Segmentation in Fundus Images Using Inception Module and Residual Connection. *IEEE Access* **2020**, *8*, 167225–167235. [CrossRef]
5. Monemian, M.; Rabbani, H. A New Method for the Localization of Hard Exudates Based on Analyzing Intensity Incremental-Decremental Trends. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–12. [CrossRef]
6. Ghalwash, A.Z.; Youssif, A.A.A.; Allam, A.M.N. Segmentation of Exudates via Color-Based K-Means Clustering and Statistical-Based Thresholding. *J. Comput. Sci.* **2017**, *13*, 524–536. [CrossRef]
7. Kurilova, V.; Goga, J.; Oravec, M.; Pavlovicova, J.; Kajan, S. Support Vector Machine and Deep-Learning Object Detection for Localisation of Hard Exudates. *Sci. Rep.* **2021**, *11*, 16045. [CrossRef] [PubMed]
8. Kumar, A. An Overview of Deep Learning Methods for Exudate Detection in Diabetic Retinopathy. *J. Eng. Sci.* **2024**, *20*, 5384–5401. [CrossRef]
9. Zhao, X.; Wang, H.; Peng, Z. Fine Segmentation of Fundus Optic Disc Based on SBF Superpixel. In *Proceedings of SPIE Medical Imaging*; SPIE: Bellingham, WA, USA, 2019; Volume 11342. [CrossRef]
10. Ashraf, M.N.; Hussain, M.; Habib, Z. Review of Various Tasks Performed in the Preprocessing Phase of a Diabetic Retinopathy Diagnosis System. *Curr. Med. Imaging* **2020**, *16*, 397–426. [CrossRef] [PubMed]
11. Xu, L.L.; Jia, B.X. Automatic White Balance Based on Gray World Method and Retinex. *Appl. Mech. Mater.* **2013**, *462–463*, 837–840. [CrossRef]
12. Yao, M.; Zhu, C. Study and Comparison on Histogram-Based Local Image Enhancement Methods. In *Proceedings of the International Conference on Image, Vision and Computing (ICIVC)*; IEEE: New York, NY, USA, 2017; pp. 309–314. [CrossRef]
13. Sharma, R.; Kamra, A. A Review on CLAHE Based Enhancement Techniques. In *Proceedings of the International Conference on Computational Intelligence and Intelligent Systems*; IEEE: New York, NY, USA, 2023; pp. 321–325. [CrossRef]
14. Priyadharsini, C. Retinal Image Enhancement Based on Color Dominance of Image. *Sci. Rep.* **2023**, *13*, 7172. [CrossRef] [PubMed]
15. Lukashchuk, B.; Shabatura, Y.V. Methodology for Evaluating Complex Object Contour Detection Accuracy in SBF-Based Image Segmentation. *Sci. Bull. UNFU* **2024**, *34*, 13–20. [CrossRef]
16. Zhou, W.; Wu, C.; Chen, D.; Wang, Z.; Yi, Y.; Du, W. A Novel Approach for Red Lesions Detection Using Superpixel Multi-Feature Classification in Color Fundus Images. In Proceedings of the Chinese Control and Decision Conference (CCDC), Chongqing, China, 28–30 May 2017; pp. 6643–6648. [CrossRef]
17. Godlevsky, L.S.; Kresyun, N.V.; Martsenyuk, V.; Shakun, K.S.; Tatarchuk, T.V.; Prybolovets, K.O.; Kalinichenko, L.F.; Karpinski, M.; Gancarczyk, T. Information System for Early Diabetic Retinopathy Diagnostics Based on Multiscale Texture Gradient Method. *Int. J. Med. Health Sci.* **2020**, *14*, 218–221.
18. Tareef, A. Superpixel-Based C-SVC for Brain Tissue Classification in MRI Scans. *Eng. Technol. Appl. Sci. Res.* **2024**, *14*, 18271–18276. [CrossRef]
19. Belém, F.; Perret, B.; Cousty, J.; Guimarães, S.J.F.; Falcão, A.X. Efficient Multiscale Object-Based Superpixel Framework. *J. Braz. Comput. Soc.* **2025**, *31*, 355–372. [CrossRef]
20. Sariera, T.M.A.; Rangarajan, L.; Amarnath, R. Detection and Classification of Hard Exudates in Retinal Images. *J. Intell. Fuzzy Syst.* **2020**, *38*, 1943–1949. [CrossRef]
21. Wu, M.; Leng, T.; de Sisternes, L.; Rubin, D.L.; Chen, Q. Segmentation of Optic Disc and Cup-to-Disc Ratio Quantification Based on OCT Scans. In *Machine Learning in Healthcare Imaging*; Springer: Singapore, 2019; pp. 193–209. [CrossRef]
22. Irshad, S.; Yin, X.; Li, L.Q.; Salman, U. Automatic Optic Disk Segmentation in Presence of Disk Blurring. In *Ophthalmic Medical Image Analysis*; Springer: Cham, Switzerland, 2016; pp. 13–23. [CrossRef]
23. Yinghua, M.; Heng, Y.; Amarnath, R.; Zeng, H. Hard Exudates Segmentation in Diabetic Retinopathy Using DiaRetDB1. *IEEE Access* **2024**, *12*, 126486–126502. [CrossRef]
24. Sheng, H.; Du, H.; Shen, X.; Wang, S.; Yu, X. Multimodal Retina Image Analysis Survey: Datasets, Tasks and Methods. In Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI), Jeju, Republic of Korea, 3–9 August 2024; pp. 10650–10659. [CrossRef] [PubMed]
25. Shivakumar, K.; Sandhya, B. Segmentation of Hard Exudates in Fundus Images to Detect Diabetic Retinopathy Using Modified U-Net. *Int. J. Sci. Res. Arch.* **2023**, *10*, 1069–1075. [CrossRef]
26. Pawlik, T.M. Textural and Statistical Feature Extraction from Segmented Hard Exudates for Diabetic Retinopathy Classification. In *Smart Innovation, Systems and Technologies*; Springer: Singapore, 2022; pp. 293–301. [CrossRef]
27. Monemian, M.; Rabbani, H. Exudate Identification in Retinal Fundus Images Using Precise Textural Verifications. *Sci. Rep.* **2023**, *13*, 2824. [CrossRef] [PubMed]

28. Lipton, Z.C.; Elkan, C.; Naryanaswamy, B. Optimal Thresholding of Classifiers to Maximize F1 Measure. In *Machine Learning and Knowledge Discovery in Databases*; Springer: Berlin/Heidelberg, Germany, 2014; Volume 8725, pp. 225–239. [CrossRef] [PubMed]
29. Qomariah, D.U.N.; Tjandrasa, H.; Fatichah, C. Exudate Segmentation for Diabetic Retinopathy Using Modified FCN-8 and Dice Loss. *Int. J. Intell. Eng. Syst.* **2022**, *15*, 508–520. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.