Journal of Computing Theories and Applications ISSN:3024-9104

Research Article

Retinal Vessel Segmentation via Morphological Refinement and Adaptive Late Fusion

Afrig Aminuddin [1,*], Mohammad Badrul Alam Miah [2], Ahmed Adil Nafea [3], and Hesmeralda Rojas Enriquez [4]

[1] Department of Information Systems, Faculty of Computer Science, Universitas Amikom Yogyakarta, Sleman, 55283, Indonesia; e-mail : afrig@amikom.ac.id
[2] Department of Information and Communication Technology (ICT), Mawlana Bhashani Science and Technology University (MBSTU), Tangail-1902, Bangladesh; e-mail : badrul.ict@mbstu.ac.bd
[3] Department of Artificial Intelligence, College of Computer Science and IT, University of Anbar, Ramadi 31001, Iraq; e-mail : ahmed.a.n@uoanbar.edu.iq
[4] Escuela Académico Profesional de Ingeniería Informática y Sistemas, Facultad de Ingeniería, Universidad Nacional Micaela Bastidas de Apurímac, Abancay, 03001, Peru; e-mail : hrojas@unamba.edu.pe
* Corresponding Author : Afrig Aminuddin

Abstract: Retinal vessel segmentation is fundamental for quantitative retinal vascular analysis; however, accurate delineation remains challenging because of nonuniform illumination, low-contrast capillaries, pathological lesions, and variations in image acquisition. This study presents a rule-based unsupervised retinal vessel segmentation framework based on decision-level adaptive late fusion, in which six complementary local adaptive thresholding methods are integrated and subsequently refined through luminance validation, hysteresis reconstruction, component cleanup, elongation filtering, and boundary smoothing. In this context, unsupervised denotes that no statistical segmentation model is trained using manual vessel annotations; instead, fixed parameters and fusion weights are determined from a small development subset, while all evaluation images remain unseen during method development. The proposed framework was evaluated on 135 independent retinal fundus images from the DRIVE, STARE, CHASE_DB1, HRF, and LES-AV datasets using an eroded field-of-view protocol. It achieved an image-weighted Dice coefficient of 0.7104 (95% bootstrap confidence interval: 0.7005–0.7205), an IoU of 0.5539, a sensitivity of 0.7544, a specificity of 0.9597, a balanced accuracy of 0.8570, and a clDice score of 0.7395. Compared with fixed majority voting, the proposed adaptive late fusion strategy significantly improved segmentation performance in 128 of 135 test images, yielding a mean Dice improvement of 0.0270 (paired Wilcoxon, Holm-adjusted $p = 7.67 \times 10^{-21}$). Although segmentation performance remains limited for complex pathological images, particularly in the HRF dataset, the proposed framework demonstrates consistent cross-dataset performance, transparent decision-making, and strong reproducibility, providing an effective alternative when interpretable, training-free retinal vessel segmentation is required.

Received: April, 30th 2026
Revised: July, 20th 2026
Accepted: July, 21st 2026
Published: July, 30th 2026

Keywords: Adaptive late fusion; Fundus imaging; Image segmentation; Local adaptive thresholding; Morphological refinement; Retinal vessel segmentation; Training-free segmentation; Unsupervised segmentation.

1. Introduction

Retinal fundus imaging is a fundamental, non-invasive diagnostic modality for the early detection and longitudinal monitoring of severe ocular and systemic diseases, including diabetic retinopathy, glaucoma, and hypertension [1]. Quantitative assessment of retinal vascular morphology, such as vessel tortuosity, branching angles, and vascular density, provides clinically important biomarkers for disease diagnosis and progression [2]. However, manual pixel-wise delineation of retinal vessels is labor-intensive, inherently subjective, and susceptible to considerable inter-observer variability [3]. These limitations motivate the development of automated retinal vessel segmentation frameworks that are accurate, robust, and reproducible.

Recent advances in supervised deep learning have substantially improved retinal vessel segmentation performance. Nevertheless, their practical deployment remains constrained by the availability of representative annotated datasets and the effects of acquisition-domain variation [4]. Supervised models require manually annotated images to optimize their parameters [5], while producing high-quality medical annotations is expensive and time-consuming [6]. Furthermore, performance may degrade when models trained on standardized datasets are applied to images acquired under different imaging conditions, including variations in spatial resolution, illumination, imaging devices, and pathological characteristics.

To reduce this dependency on annotated training data, deterministic rule-based approaches remain an attractive alternative, particularly those based on local adaptive thresholding [7]. However, individual thresholding algorithms exhibit inherent trade-offs between preserving low-contrast capillaries and suppressing false-positive responses, with performance varying across imaging conditions and retinal pathologies [8]–[10]. Existing multi-threshold approaches attempt to alleviate these limitations through algorithm fusion; however, most rely on static majority voting strategies that remain vulnerable to correlated errors and image-dependent variability.

To address these challenges, this study proposes a rule-based unsupervised retinal vessel segmentation framework based on morphological refinement and adaptive decision-level late fusion. Instead of relying on a fixed consensus strategy, the proposed framework combines six complementary local adaptive thresholding methods to generate multiple binary vessel candidates, which are subsequently refined through luminance validation, hysteresis reconstruction, component cleanup, elongation filtering, and boundary smoothing. In this study, unsupervised indicates that no statistical segmentation model is trained using manual vessel annotations. Instead, a small development subset is used solely to determine fixed algorithm parameters and fusion weights, while all reported evaluation images remain completely unseen throughout method development. The primary contributions of this work are summarized as follows:

- An adaptive decision-level late fusion strategy that dynamically adjusts voting strictness according to the global agreement among multiple thresholding outputs for each retinal image.
- A morphology-guided refinement pipeline that integrates local luminance validation and elongation-based filtering to suppress non-vascular structures while preserving vessel continuity.
- A leakage-controlled experimental evaluation on 135 held-out retinal fundus images from five public datasets, including confidence interval estimation, paired statistical testing, topology-aware and boundary-based evaluation metrics, stage-wise ablation analysis, and pathology-stratified performance assessment.

The remainder of this paper is organized as follows. Section 2 reviews existing retinal vessel segmentation methods and identifies the limitations of supervised, rule-based, unsupervised, and decision-level fusion approaches. Section 3 describes the proposed framework, including preprocessing, multi-variant vessel extraction, and adaptive decision-level late fusion. Section 4 presents quantitative and qualitative evaluations, ablation studies, comparisons with constituent and published methods, pathology-stratified analyses, and a discussion of the proposed framework's limitations. Finally, Section 5 summarizes the main findings, discusses future research directions, and concludes the paper.

2. Literature Review

This section reviews recent advances in retinal vessel segmentation to establish the methodological context of this study. It examines the strengths and limitations of prevailing supervised and deterministic rule-based approaches, identifies the unresolved trade-offs associated with existing local thresholding techniques, and motivates the development of the proposed adaptive decision-level late fusion framework.

2.1. Supervised Segmentation

Recent progress in retinal vessel segmentation has been driven predominantly by supervised deep learning architectures, particularly convolutional neural networks (CNNs) and U-Net variants [11]. By learning hierarchical feature representations directly from annotated

data, these models have achieved excellent performance on standardized retinal image benchmarks [12]. Nevertheless, their deployment in routine clinical practice remains constrained by annotation availability, operational cost, and acquisition-domain variability [13].

Supervised models require representative pixel-level annotations to optimize network parameters [14]. In ophthalmic imaging, generating such annotations is labor-intensive, expensive, and inherently affected by inter-observer variability [6]. Moreover, segmentation performance may deteriorate when models trained on one imaging domain are applied to external datasets acquired under different conditions [15]. Variations in native image resolution, illumination characteristics, imaging devices, and pathological manifestations can substantially affect generalization performance, as observed across datasets such as DRIVE, HRF, and CHASE_DB1 [16]. In addition, the limited interpretability of many deep learning models may reduce diagnostic transparency in clinical applications [17]. Consequently, deterministic rule-based segmentation methods remain an attractive alternative when representative annotations are unavailable, or model retraining is impractical.

2.2. Unsupervised Thresholding

To overcome the dependency on annotated training data, deterministic rule-based segmentation methods employ mathematical heuristics to distinguish retinal vessels from the surrounding background [18]. Because retinal fundus images often exhibit pronounced illumination gradients, central light reflexes, and local contrast variation, conventional global thresholding techniques, such as Otsu's method, generally produce unsatisfactory segmentation results [19]. Consequently, most classical approaches rely on local adaptive thresholding, where the threshold for each pixel is estimated from statistics computed within its local neighborhood [20].

The six thresholding methods considered in this study adopt different statistical formulations. Niblack, Sauvola, Wolf, and Phansalkar estimate thresholds from combinations of local mean intensity and local dispersion, whereas Bradley compares each pixel with its local mean, and Bernsen derives thresholds from local intensity extrema and contrast measurements [21]. Owing to these different formulations, segmentation performance varies with algorithm parameters, illumination conditions, pathological structures, and vessel caliber. Consequently, no single thresholding method consistently achieves both high vessel sensitivity and effective suppression of false-positive responses across diverse retinal images [22].

Thresholding methods that aggressively preserve weak local contrast generally improve the detection of fine capillaries but are also more susceptible to background noise and lesion-induced false positives [23]. Conversely, more conservative thresholding strategies suppress noise more effectively but frequently produce fragmented vessel networks and reduced continuity [24], [25]. Although morphological post-processing can partially alleviate these limitations, it cannot completely eliminate the inherent trade-offs associated with individual thresholding formulations. These complementary strengths and weaknesses provide the primary motivation for integrating multiple thresholding outputs within a unified decision-level framework.

2.3. Fusion Strategies

To balance the complementary strengths and weaknesses of individual thresholding methods, numerous retinal vessel segmentation studies have explored multi-method fusion strategies [26]. Existing approaches commonly combine multiple binary vessel masks using logical operations, such as Boolean AND/OR, or simple majority voting to generate a consensus segmentation [27]. Although these decision-level fusion strategies generally improve robustness compared with a single thresholding method, they remain susceptible to correlated errors among the constituent algorithms [28]. When several thresholding methods respond similarly to bright lesions, peripheral shadows, or other non-vascular structures, majority voting may reinforce rather than suppress these shared false-positive responses [29]. Moreover, fixed voting thresholds assume consistent image quality and therefore cannot adequately accommodate the substantial variation in illumination, contrast, and signal-to-noise ratio observed across retinal fundus images.

These limitations motivate a more flexible deterministic fusion strategy that considers not only the number of votes but also the consistency of the segmentation evidence and the structural characteristics of retinal vessels [30]. Accordingly, the proposed framework

incorporates image-level adaptive vote selection, local luminance validation, and morphology-guided refinement while preserving fully deterministic and auditable decision rules.

In parallel with the development of rule-based approaches, recent retinal vessel segmentation research has increasingly adopted transformer-based architectures, domain-generalization techniques, self-supervised learning, and lightweight neural networks. Transformer and attention mechanisms enhance long-range contextual representation, whereas domain-generalization and self-supervised methods aim to improve robustness across imaging domains and reduce annotation dependency. Lightweight architectures reduce computational complexity while maintaining competitive segmentation performance. Despite these advances, all of these methods remain dependent on learned feature representations and parameter optimization. Consequently, they address a different operational scenario from the deterministic framework proposed in this study and are included here to establish methodological context rather than to serve as protocol-equivalent baselines.

Table 1 summarizes representative retinal vessel segmentation methods according to their underlying methodology, fusion or representation strategy, dataset, training requirement, and principal limitation. Unlike previous studies, the proposed framework combines multiple local adaptive thresholding methods through adaptive decision-level late fusion followed by deterministic morphological refinement, thereby maintaining complete interpretability while avoiding statistical model fitting.

Table 1. Methodological comparison of representative retinal vessel segmentation methods.

Study	Methodology	Fusion / Representation Strategy	Dataset(s)	Learning Strategy	Primary Limitation
Nandy and Banerjee [21]	Local adaptive thresholding	Independent Niblack and Sauvola thresholding	DRIVE	Fixed empirical parameters	Single-threshold formulation with limited adaptability
Zhou et al. [31]	Line detector + Hidden Markov Model	Thick- and thin-vessel integration	DRIVE, STARE	Parameterized optimization	Sensitive to acquisition variability
Mahapatra et al. [32]	Frangi enhancement + spatial FCM	Filter enhancement and clustering	DRIVE, STARE, HRF	Parameter optimization	Dataset-dependent parameter tuning
Khan et al. [33]	Multiscale directional filtering	Separate processing of thick and thin vessels	DRIVE, STARE, CHASE_DB1	Fixed empirical parameters	Multiple manually configured filters
Qu et al. [34]	Two-path CNN	Learned multiscale feature representation	DRIVE, STARE, CHASE_DB1	Supervised learning	Annotation and domain dependence
Hu et al. [35]	Domain generalization	Domain-invariant feature learning	Multiple datasets	Supervised domain generalization	Requires source-domain optimization
Tu et al. [36]; Tong et al. [37]	Self-supervised and lightweight networks	Pretrained representation and transformer features	DRIVE, STARE, CHASE_DB1	Self-supervised pretraining / fine-tuning	Dependence on learned representations
Proposed	Local adaptive thresholding + morphology	Adaptive decision-level late fusion	DRIVE, STARE, CHASE_DB1, HRF, LES-AV	Fixed parameters selected from development subset	Sensitivity to pathology and fixed parameter configuration

Because the proposed framework combines completed binary segmentation masks, the integration process is classified as decision-level (late) fusion rather than feature-level fusion. Furthermore, the term adaptive refers exclusively to the deterministic adjustment of voting thresholds according to image-level agreement among the constituent segmentations. It does not imply online parameter optimization, statistical learning, or adaptive model training.

3. Proposed Method

This section describes the proposed rule-based unsupervised retinal vessel segmentation framework. The framework does not train or fit a statistical segmentation model. Instead, manual vessel annotations from the development subset were used only to determine fixed algorithm parameters, morphological variants, and fusion weights. After this development stage, all segmentation procedures and decision rules remain fixed throughout evaluation.

3.1. General Framework

The overall workflow of the proposed framework is illustrated in Figure 1. First, all retinal images, reference annotations, and field-of-view (FOV) masks are spatially standardized to establish a common processing scale. The valid FOV is then eroded to suppress peripheral imaging artifacts before segmentation. For each of the six local adaptive thresholding methods, the standardized green channel is smoothed using its selected Gaussian scale, enhanced using Contrast-Limited Adaptive Histogram Equalization (CLAHE), and converted into an initial binary vessel mask. The corresponding enhanced image is subsequently used to estimate an optic-disc candidate and generate three morphological variants (Base, Traced, and Open). The selected variants are integrated through adaptive decision-level late fusion, followed by luminance validation, hysteresis reconstruction, component cleanup, elongation filtering, and boundary smoothing. Finally, the refined vessel mask is restricted to the valid FOV and evaluated against the aligned reference annotation.

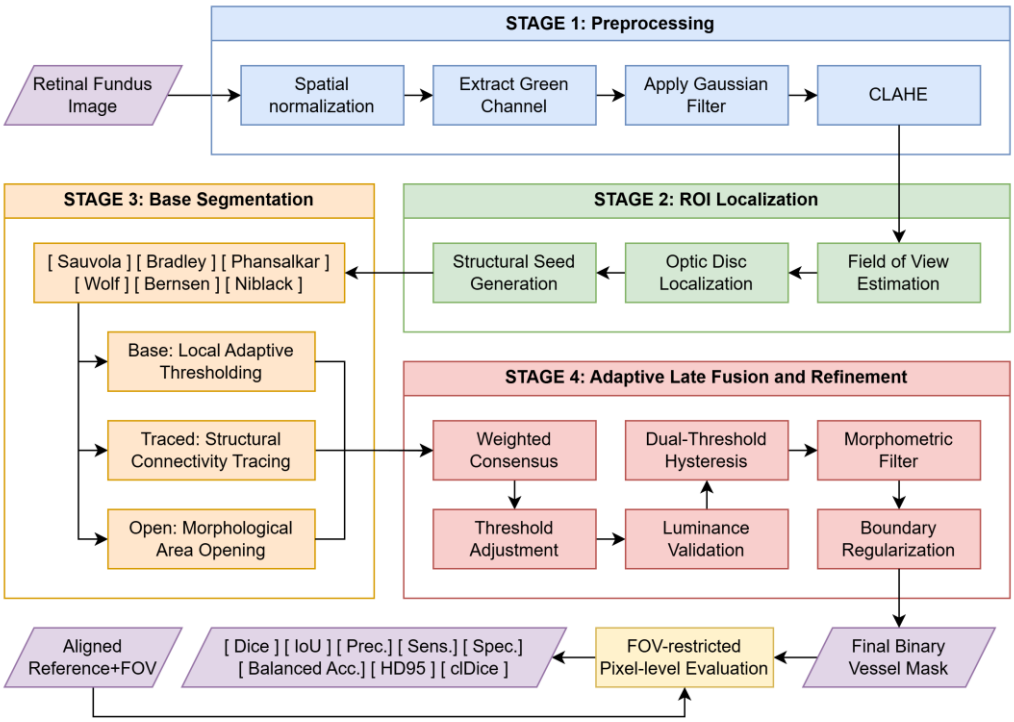

Figure 1. Overall workflow of the proposed rule-based unsupervised retinal vessel segmentation framework.

3.2. Preprocessing

The preprocessing stage standardizes image geometry and enhances vascular contrast before segmentation. It consists of four sequential operations: spatial normalization, green-channel extraction, Gaussian smoothing, and local contrast enhancement.

1. Spatial Dimension Normalization

Retinal datasets possess varying native resolutions (e.g., 565 × 584 in DRIVE versus 700 × 605 in STARE). The retinal field is estimated from its boundary, cropped, scaled to a retinal diameter of 1,984 pixels, and centered on a 2,048 × 2,048 canvas. This normalization makes subsequent pixel-scale parameters approximately comparable; it is not evidence of resolution invariance. Intensity images are resized using bicubic interpolation. Manual references and FOV masks undergo the identical crop, translation, and scale transformation using nearest-neighbor interpolation to preserve binary labels. Filename, image dimensions, and overlay alignment are checked before evaluation. This transformation is formulated as:

$$I_{\text{scaled}} = \mathcal{P}_{2048}\left(\mathcal{B}_s\left(I_{\text{crop}}\right)\right) \tag{1}$$

$$s = \frac{1984}{d_{\text{retina}}} \tag{2}$$

where I_{crop} is the retinal bounding crop, d_{retina} is its estimated retinal diameter, $\mathcal{B}_s$ denotes bicubic scaling by factor s, and $\mathcal{P}_{2048}$ centers the scaled retinal field on a 2048 × 2048 canvas. The identical crop, scale, and centering transformation is applied to the reference annotation and FOV mask using nearest-neighbor interpolation rather than $\mathcal{B}_s$.

2. Spectral Component Isolation:

The green spectral component exhibits the highest optical absorption rate by retinal hemoglobin. Isolating this channel maximizes the visual contrast between the vascular network and the surrounding retinal background, which is essential for accurate intensity-based binarization. The extraction is defined as:

$$I_g(x, y) = I_{scaled}(x, y, 2) \tag{3}$$

where I_g represents the extracted single-channel intensity matrix, and the index 2 denotes the extraction of the green channel from the RGB color space.

3. Stochastic Noise Suppression:

Raw fundus captures frequently contain high-frequency stochastic sensor noise that local thresholding algorithms often misclassify as micro-vessels. A Gaussian filter is applied to attenuate these microscopic artifacts without severely degrading the boundary gradients of thin capillaries, expressed as:

$$I^k_{\text{smooth}}(x, y) = \sum_{u,v} I_g\,(x - u, y - v) G_{\sigma_k}(u, v) \tag{4}$$

$$G_{\sigma_k}(u, v) = \frac{1}{2\pi\sigma_k^2} \exp\left[-\frac{u^2 + v^2}{2\sigma_k^2}\right] \tag{5}$$

where I^k_{smooth} is the image smoothed for threshold family k, G_{σ_k} is the two-dimensional Gaussian kernel, and σ_k is selected separately for each family using the development subset (Table 2).

4. Local Histogram Equalization:

Global thresholding fails in retinal imagery due to the central light reflex and severe illumination gradients, such as peripheral vignetting. Contrast-Limited Adaptive Histogram Equalization (CLAHE) is utilized to normalize contrast within discrete localized tiles, effectively highlighting fine vessels in under-illuminated regions while preventing the over-amplification of background noise. This localized enhancement is given by:

$$I^k_{\text{clahe}} = \mathcal{C}_{\text{CLAHE}}\left(I^k_{\text{smooth}};\, N_x \times N_y,\, C_{\text{limit}}\right) \tag{6}$$

where I^k_{clahe} is the enhanced image for threshold family k, $\mathcal{C}_{\text{CLAHE}}$ denotes contrast-limited adaptive histogram equalization, $N_x \times N_y = 8 \times 8$ is the tile grid, and $C_{\text{limit}} = 0.01$ is the clip limit. Each method-specific enhanced image I^k_{clahe} is used for its corresponding optic-disc estimate, morphological variants, and local threshold calculation.

3.3. ROI Localization

The ROI localization phase delineates the valid retinal field and estimates an optic-disc candidate. These spatial constraints provide the valid processing region and the anchor required for subsequent morphological reconstruction. The procedure is executed through the following sequential steps:

1. Field of View Estimation:

Fundus images are strictly confined to the retinal projection within a circular camera aperture, often leaving dark peripheral regions that produce false-positive gradients during local binarization. A binary Field of View (FOV) mask is established and subsequently morphologically eroded to contract the boundary inward, thereby safely excluding these severe peripheral illumination artifacts. This operation is defined as:

$$M_{FOV} = M_{initial} \ominus S_{disk}(r_{erode}) \tag{7}$$

where M_{FOV} is the refined valid region mask, $M_{initial}$ is the initial binary aperture mask, $\ominus$ denotes the morphological erosion operator, and $S_{disk}(r_{erode})$ is a two-dimensional disk-shaped structuring element with a radius $r_{erode} = 8$ pixels.

2. Optic Disc Localization:

For each threshold family, an optic-disc candidate is estimated as the maximum of a strongly smoothed version of the corresponding enhanced image inside the valid FOV. This location provides the anchor used to construct that family's Traced variant:

$$\left(x_{\text{od}}^{k}, y_{\text{od}}^{k}\right) = \underset{(x,y)\in\Omega_{\text{FOV}}}{\arg\max} \left(G_{\sigma_{\text{od}}} * I_{\text{clahe}}^{k}\right)(x,y) \tag{8}$$

where $\left(x_{\text{od}}^{k}, y_{\text{od}}^{k}\right)$ is the candidate location for family k, Ω_{FOV} is the valid FOV, and $G_{\sigma_{\text{od}}}$ is a Gaussian kernel with $\sigma_{\text{od}} = 35$ pixels. This is an intensity-based candidate location rather than a separately validated optic-disc segmentation.

3. Structural Seed Generation:

A method-specific circular seed region is constructed around the candidate location for morphological reconstruction:

$$M_{\text{seed}}^{k}(x,y) = \mathbf{1}\left[\left(x - x_{\text{od}}^{k}\right)^2 + \left(y - y_{\text{od}}^{k}\right)^2 \le r_{\text{od}}^2\right] \tag{9}$$

where M_{seed}^{k} is the binary anchor mask and $r_{\text{od}} = 160$ pixels. Spatial standardization makes this fixed radius approximately comparable across datasets; it does not establish scale invariance.

3.4. Base Segmentation

The base segmentation phase translates the enhanced continuous-tone image into discrete binary representations of the vascular network. Because no single local thresholding metric performs optimally across all vessel thicknesses and contrast variations, this framework executes a heterogeneous array of binarization models in parallel. To further isolate true vascular structures from algorithmic noise, each thresholded output is subjected to specific morphological filtering, producing three distinct variants (Base, Traced, and Open) per algorithm. The procedure is executed through the following sequential steps:

1. Local Adaptive Thresholding:

Varying pathologies, central light reflexes, and uneven background intensities require threshold surfaces computed from local pixel neighborhoods rather than global statistics. Six distinct local thresholding functions—Sauvola, Bradley, Phansalkar, Wolf, Bernsen, and Niblack—are applied independently to the preprocessed image to generate an initial set of raw binary masks. This parallel binarization is generalized as:

$$M_{\text{base}}^{k}(x,y) = \Phi_k\left(I_{\text{clahe}}^{k}(x,y); w_k, \theta_k\right) \cap M_{\text{FOV}}(x,y) \tag{10}$$

where M_{base}^{k} is the FOV-restricted binary output of threshold family k, Φ_k is the corresponding threshold function, and w_k and θ_k are its development-selected window size and principal parameter, respectively.

2. Structural Connectivity Tracing:

Local thresholding inherently produces isolated islands of false-positive noise in the retinal background. Morphological reconstruction by dilation is employed to preserve only the contiguous vascular tree, discarding independent noise artifacts. The predefined optic disc region serves as the intersection seed, ensuring that only structures physically connected to the central anatomical anchor are retained:

$$M_{\text{trace}}^{k} = \rho_{M_{\text{base}}^{k}}\left(M_{\text{seed}}^{k} \cap M_{\text{base}}^{k}\right) \tag{11}$$

where M_{trace}^{k} is the reconstructed Traced variant and $\rho_{M_{\text{base}}^{k}}(\cdot)$ denotes morphological reconstruction by dilation under the constraining mask M_{base}^{k}.

3. Morphological Area Opening:

While the tracing operation guarantees macro-connectivity, some algorithms generate highly sensitive boundary noise or "hairy" artifacts that are weakly attached to the main vessels. A targeted neck-cut sequence is applied via erosion to detach these weakly connected pixels, followed by an area-opening filter to eliminate them. The vessels are subsequently dilated back to their original morphological boundaries and intersected with the base mask to prevent artificial expansion:

$$M_{eroded}^{k} = M_{base}^{k} \ominus S_{disk}(r_1) \tag{12}$$

$$M_{open}^{k} = \left(\gamma_{area}\left(M_{eroded}^{k}, A_{min}\right) \oplus S_{disk}(r_1)\right) \cap M_{base}^{k} \tag{13}$$

where M_{open}^{k} is the cleaned area-open mask, $\ominus$ and $\oplus$ are the morphological erosion and dilation operators, respectively. $S_{disk}(r_1)$ is a disk structuring element with a radius of 1 pixel, and γ_{area} removes connected components with fewer than $A_{min} = 50$ pixels.

This multi-variant approach yields a comprehensive set of 18 distinct binary representations of the vascular network, providing the necessary methodological and morphological diversity for the subsequent adaptive ensemble fusion phase. Parameter selection retained one variant from each family for fusion. The complete development-selected configuration is reported in Table 2.

Table 2. Threshold configurations selected using the development subset

Threshold method	Selected variant	Window	Principal parameter	Gaussian σ	Development Dice
Sauvola	Base	71	k=0.15	3.00	0.7265
Bradley	Open	63	t=10	3.00	0.7260
Phansalkar	Base	79	k=0.20	3.00	0.6896
Wolf	Base	71	k=0.15	2.25	0.7251
Bernsen	Traced	15	contrast=0.09	3.00	0.7071
Niblack	Traced	63	k=−0.40	1.50	0.6980

Phansalkar constants were fixed at $R = 0.5$, $p = 2$, and $q = 10$. Common settings were CLAHE with an 8 × 8 tile grid and clip limit 0.01, FOV erosion radius 8 pixels, optic-disc smoothing $\sigma = 35$ pixels, and structural seed radius 160 pixels.

3.5. Ensemble Fusion

The ensemble-fusion stage combines the development-selected binary representations into a single binary vascular mask. Because simple majority voting can preserve correlated errors among component masks, the proposed framework employs weighted late fusion with deterministic threshold adjustment. Global voting coherence, a local luminance constraint, and sequential geometric refinement are used to suppress a proportion of the non-vascular responses retained by the component masks. The procedure is executed through the following sequential steps:

1. Weighted Voting Logic:

The development-selected morphological variant from each threshold family contributes to the weighted consensus surface. Fixed heuristic integer weights, established during development, specify the relative contribution of each selected variant. These weights are neither statistically fitted coefficients nor quantities recalculated during inference. The weighted consensus surface is defined as:

$$V(x,y) = \sum_{k=1}^{6} w_k \cdot M_{variant}^{k}(x,y) \tag{14}$$

where $V(x,y)$ represents the accumulated voting score at each spatial coordinate, M_{variant}^{k} is the selected morphological mask (Base, Traced, or Open) for threshold family k, and w_k denotes its fixed weight. The fixed integer weights are Bradley-Open = 3, Sauvola-Base = 2,

Wolf-Base = 2, and Phansalkar-Base, Bernsen-Traced, and Niblack-Traced = 1 each, yielding a maximum vote of 10. These are heuristic ranks based on development performance, not learned coefficients.

2. Dynamic Consensus Evaluation:

Fusion strictness is selected deterministically according to the global coherence of the weighted voting surface. The agreement ratio (AR) measures the proportion of voting-positive pixels that also satisfy the high-agreement bound within the valid FOV. This ratio determines the thresholds for the high-confidence structural seeds (τ_{seed}) and the permissive reconstruction network (τ_{net}):

$$AR = \frac{\sum_{(x,y)\in\Omega_{\text{FOV}}} \mathbf{1}\left[V(x,y) \geq \tau_{\text{high}}\right]}{\sum_{(x,y)\in\Omega_{\text{FOV}}} \mathbf{1}\left[V(x,y) \geq \tau_{\text{low}}\right] + \varepsilon} \tag{15}$$

where Ω_{FOV} denotes the valid field of view, $\mathbf{1}[\cdot]$ is the indicator function, and ε is machine epsilon, included to prevent division by zero. The static voting-score bounds are $\tau_{\text{high}} = 7$ and $\tau_{\text{low}} = 1$. Thus, AR is the number of pixels inside the FOV with a weighted score of at least 7 divided by the number receiving a weighted score of at least 1. The default reconstruction thresholds are $\tau_{\text{seed}} = 7$ and $\tau_{\text{net}} = 4$, for a maximum weighted score of 10. When $AR > 0.45$, they are relaxed to $\tau_{\text{seed}} = 6$ and $\tau_{\text{net}} = 3$; when $AR < 0.20$, they are increased to $\tau_{\text{seed}} = 8$ and $\tau_{\text{net}} = 5$. Otherwise, the default values are retained. These rules were fixed during development; therefore, adaptation denotes deterministic threshold selection rather than online learning.

3. Background Luminance Validation:

A separate validation image is formed from the standardized green channel using Gaussian smoothing with $\sigma_{\text{lum}} = 2$ pixels and CLAHE. A 51-pixel box filter estimates its local background, and pixels are retained when they are darker than that background by the fixed offset δ:

$$I_{\text{lum}} = \mathcal{C}_{\text{CLAHE}}\left(G_{\sigma_{\text{lum}}} * I_g; 8 \times 8, 0.01\right) \tag{16}$$

$$I_{\text{bg}} = \mathcal{B}_{51}(I_{\text{lum}}) \tag{17}$$

$$M_{\text{lum}}(x,y) = \mathbf{1}\left[I_{\text{lum}}(x,y) < I_{\text{bg}}(x,y) - \delta\right] \tag{18}$$

where $\mathcal{B}_{51}$ denotes a 51 × 51 box filter with replicated boundary padding, $\sigma_{\text{lum}} = 2$ pixels, and $\delta = 0.002$. The validation image I_{lum} is separate from the method-specific images I^{k}_{clahe}.

4. Dual-Threshold Hysteresis Integration:

To construct the final contiguous network, high-confidence pixels are utilized as structural seeds and morphologically reconstructed into the broader, luminance-validated voting network. This dual-threshold hysteresis recovers thin, low-contrast capillaries that lack sufficient independent voting weight, provided they physically connect to the primary vascular tree:

$$M_{\text{net}}(x,y) = \mathbf{1}[V(x,y) \geq \tau_{\text{net}}] \cap M_{\text{lum}}(x,y) \cap M_{\text{FOV}}(x,y) \tag{19}$$

$$M_{\text{fused}} = \rho_{M_{\text{net}}}(\mathbf{1}[V(x,y) \geq \tau_{\text{seed}}] \cap M_{\text{net}}) \tag{20}$$

where M_{net} is the luminance-validated, FOV-restricted permissive network; M_{fused} is the reconstructed fusion mask; and $\rho_{M_{\text{net}}}(\cdot)$ denotes morphological reconstruction under M_{net}.

5. Morphometric Elongation Filtering:

Pathological lesions and residual noise clusters frequently present as compact morphological islands, whereas genuine vessels are elongated and tubular. A geometric heuristic is applied to systematically eliminate non-vascular structures by evaluating the Fill Ratio (FR) of every independent connected component C_i:

$$FR(C_i) = \frac{A(C_i)}{\pi\left[L_{\text{major}}(C_i)/2\right]^2 + \varepsilon}, \qquad C_i \text{ is retained if } FR(C_i) \leq 0.20 \tag{21}$$

where $A(C_i)$ is the component area in pixels, $L_{\text{major}}(C_i)$ is its major-axis length, and ε prevents division by zero. Components with $FR(C_i) > 0.20$ are removed. Before elongation filtering, connected components smaller than 50 pixels are removed, and holes up to 20 pixels are filled. The fill ratio is computed as component area divided by the area of a circle whose diameter is the component major-axis length.

6. Boundary Regularization:

Morphological Boolean operations and discrete pixel voting inherently introduce artificial, jagged boundary artifacts. To reflect the natural, continuous morphology of retinal vessels, a final spatial regularization is performed via slight Gaussian smoothing and binarization:

$$M_{\text{final}}(x,y) = \mathbf{1}\left[\left(G_{\sigma_b} * M_{\text{filtered}}\right)(x,y) > 0.5\right] \cap M_{\text{FOV}}(x,y) \tag{22}$$

where $\sigma_b = 1$, M_{filtered} is the mask after cleanup and elongation filtering, G_{σ_b} is the Gaussian smoothing kernel, and the final FOV intersection reproduces the implementation's last masking operation. Smoothing uses $\sigma = 1$ pixel followed by binarization at 0.5.

4. Results and Discussion

This section evaluates the proposed rule-based unsupervised retinal vessel segmentation framework using a strictly held-out protocol. It first describes the experimental setting, evaluation datasets, region of interest, and performance measures. The quantitative analysis then presents dataset-specific and pooled performance, followed by controlled ablation studies, comparisons with the selected constituent threshold outputs and published methods, pathology-stratified analysis, and qualitative assessment together with the principal methodological limitations.

4.1. Experimental Setup

All experiments were conducted in MATLAB on a workstation equipped with an AMD Ryzen 7 5700U processor and 32 GB of memory. Independent images were processed in parallel. The DRIVE training partition [38] (20 images) was reserved exclusively for method development. Eight deterministically selected and evenly spaced images were used to determine the reported parameter configuration, whereas the remaining 12 DRIVE training images were excluded from both parameter selection and performance evaluation. The held-out evaluation comprised 135 previously unseen retinal fundus images from DRIVE test (20), STARE [39] (20), CHASE_DB1 [40] (28), HRF [41] (45), and LES-AV [42] (22). No held-out reference annotation was used to determine threshold parameters, morphological variants, fusion weights, or refinement settings.

Performance was evaluated within the corresponding field of view after applying an eight-pixel erosion. Dice, IoU, precision, sensitivity, specificity, balanced accuracy, accuracy, HD95, and clDice [43] were calculated. Balanced accuracy was defined as the arithmetic mean of sensitivity and specificity. Because the framework produces binary segmentation masks rather than calibrated probability maps, ROC AUC was not reported. Unless otherwise stated, pooled results are presented as the image-weighted mean across all 135 evaluation images to account for the unequal number of samples contributed by each dataset. Ninety-five-percent confidence intervals were estimated using image-level bootstrap resampling. Statistical significance for paired ablation experiments was assessed using two-sided Wilcoxon signed-rank tests with Holm adjustment across the nine reported evaluation metrics for each specified pairwise comparison.

4.2. Held-Out Performance

Table 3 summarizes the held-out segmentation performance of the proposed framework on each evaluation dataset together with the pooled results across all 135 retinal fundus images. The proposed framework achieved an overall accuracy of 0.9380. The corresponding 95% bootstrap confidence intervals were 0.7005–0.7205 for Dice, 0.5424–0.5655 for IoU, 0.6712–0.7118 for precision, 0.7405–0.7679 for sensitivity, 0.9564–0.9628 for specificity,

0.8510–0.8626 for balanced accuracy, 58.20–68.20 pixels for HD95, and 0.7291–0.7489 for clDice. For reference, the dataset-macro Dice score was 0.7185, although the image-weighted average is considered the primary summary because the datasets contain unequal numbers of evaluation images.

Across the five datasets, Dice exceeded 0.70 for DRIVE, STARE, CHASE_DB1, and LES-AV, whereas HRF produced the lowest overlap (0.6763). DRIVE achieved the highest precision but the lowest sensitivity, indicating a relatively conservative segmentation that favors false-positive suppression. STARE obtained the highest clDice, suggesting the strongest preservation of vascular connectivity, while LES-AV achieved the highest Dice despite exhibiting the largest HD95, demonstrating that favorable region overlap does not necessarily imply optimal boundary localization. These observations highlight the complementary nature of overlap, boundary-distance, and topology-aware metrics and support the use of multiple evaluation criteria rather than relying solely on accuracy or specificity in a highly background-dominated segmentation task.

Table 3. Held-out segmentation performance of the proposed framework across five retinal fundus datasets.

Dataset (n)	Dice	IoU	Precision	Sensitivity	Specificity	Balanced Accuracy	HD95 (px)	clDice
DRIVE (20)	0.7382	0.5859	0.8309	0.6705	0.9793	0.8249	67.86	0.7398
STARE (20)	0.7286	0.5759	0.7260	0.7580	0.9630	0.8605	60.44	0.7954
CHASE_DB1 (28)	0.7068	0.5479	0.6501	0.7824	0.9517	0.8670	55.44	0.7277
HRF (45)	0.6763	0.5138	0.6003	0.7847	0.9470	0.8659	47.53	0.7224
LES-AV (22)	0.7428	0.5946	0.7739	0.7300	0.9748	0.8524	101.16	0.7381
All images (135)	0.7104	0.5539	0.6917	0.7544	0.9597	0.8570	62.83	0.7395

Based on these results, the proposed framework demonstrated consistent segmentation performance across five datasets acquired under different imaging conditions without dataset-specific parameter adaptation or model retraining. Although performance decreased on the more challenging HRF dataset, the reduction remained moderate relative to the overall pooled results, indicating that the fixed rule-based pipeline generalized reasonably well despite substantial differences in image resolution, illumination characteristics, and pathological complexity. This behavior is consistent with the objective of the proposed framework, namely to provide an interpretable and reproducible segmentation strategy that maintains stable performance across heterogeneous retinal datasets rather than optimizing specifically for a single benchmark.

4.3. Controlled Ablation

To quantify the contribution of each processing stage, a controlled ablation study was conducted by sequentially enabling the proposed fusion and refinement components. The image-weighted results are summarized in Table 4.

Table 4. Image-weighted performance of the controlled fusion and refinement stages.

Configuration	Dice	IoU	Precision	Sensitivity	Specificity	Balanced Accuracy	HD95	clDice
Fixed majority	0.6834	0.5236	0.6405	0.7684	0.9463	0.8574	52.68	0.7063
Fixed weighted	0.6757	0.5151	0.6204	0.7789	0.9406	0.8598	52.95	0.7107
Adaptive weighted	0.6722	0.5113	0.6113	0.7848	0.9375	0.8611	53.43	0.7075
+ luminance validation	0.6722	0.5112	0.6124	0.7824	0.9381	0.8602	53.40	0.7074
+ hysteresis	0.6753	0.5147	0.6191	0.7792	0.9401	0.8597	55.47	0.7143
+ cleanup	0.6754	0.5149	0.6194	0.7792	0.9402	0.8597	55.78	0.7160
+ elongation	0.7104	0.5539	0.6916	0.7545	0.9596	0.8571	62.77	0.7342
+ smoothing (Complete)	0.7104	0.5539	0.6917	0.7544	0.9597	0.8570	62.83	0.7395

As shown in Table 4, replacing fixed majority voting with fixed weighting or deterministic adaptive threshold selection alone did not improve Dice performance. Likewise,

luminance validation produced only negligible changes. Hysteresis reconstruction provided a modest improvement in overlap and centerline preservation, while the subsequent cleanup stage yielded only a marginal additional benefit. In contrast, elongation filtering produced the largest performance gain. Relative to the cleanup stage, Dice increased by 0.0349 and precision by 0.0722, whereas sensitivity decreased by 0.0247 and HD95 increased by 6.99 pixels. These results indicate that suppressing compact non-vascular structures substantially reduces false-positive responses but may also remove short vessel fragments.

Boundary smoothing produced almost no change in regional overlap compared with elongation filtering alone (raw $p = 0.194$; Holm-adjusted $p = 0.266$), but significantly improved centerline agreement, increasing clDice by 0.0053 (median paired change = +0.00378; raw $p = 1.67 \times 10^{-23}$; Holm-adjusted $p = 1.50 \times 10^{-22}$). Relative to fixed majority voting, the complete framework improved Dice in 128 of 135 evaluation images, with a mean Dice increase of 0.0270 (median = +0.01836; raw $p = 1.28 \times 10^{-21}$; Holm-adjusted $p = 7.67 \times 10^{-21}$). Improvements were observed in Dice, IoU, precision, specificity, and clDice, whereas sensitivity and HD95 deteriorated slightly. Balanced accuracy showed no statistically significant difference (median = −0.0001; raw and Holm-adjusted $p = 0.612$).

The selected standalone thresholding methods achieved global Dice scores of 0.6842 (Sauvola-Base), 0.6726 (Bradley-Open), 0.6718 (Wolf-Base), 0.6088 (Niblack-Traced), 0.5840 (Bernsen-Traced), and 0.5319 (Phansalkar-Base). The complete framework exceeded every selected constituent segmentation in Dice, although Sauvola-Base retained a lower HD95 (52.90 pixels). The per-dataset Dice results for each ablation stage are summarized in Table 5.

Table 5. Per-dataset Dice scores for the controlled fusion and refinement stages.

Configuration	DRIVE	STARE	CHASE_DB1	HRF	LES-AV
Fixed majority	0.7271	0.6967	0.6519	0.6573	0.7252
Fixed weighted	0.7250	0.6961	0.6350	0.6482	0.7201
Adaptive weighted	0.7221	0.6927	0.6294	0.6447	0.7190
+ luminance validation	0.7212	0.6930	0.6273	0.6462	0.7189
+ hysteresis	0.7238	0.6945	0.6358	0.6477	0.7204
+ cleanup	0.7238	0.6945	0.6364	0.6477	0.7206
+ elongation	0.7382	0.7284	0.7066	0.6765	0.7427
+ smoothing (Complete)	0.7382	0.7286	0.7068	0.6763	0.7428

The per-dataset analysis demonstrates a consistent pattern across all five evaluation datasets. Adaptive threshold selection and luminance validation alone produced little improvement in regional overlap, whereas elongation filtering consistently generated the principal increase in Dice regardless of dataset characteristics. Boundary smoothing had minimal influence on Dice but consistently improved centerline preservation. Overall, these observations indicate that the primary benefit of the proposed framework arises from morphology-guided refinement rather than from adaptive voting alone, emphasizing the importance of structural false-positive suppression in achieving robust vessel segmentation across heterogeneous retinal images.

4.4. Benchmarking Analysis

To provide performance context, Table 6 compares the proposed framework with representative retinal vessel segmentation studies evaluated on shared datasets. The proposed results correspond to the held-out evaluation protocol used in this study. Published binary F1 scores are reported as Dice whenever the conventional binary definition was employed, whereas unavailable values are indicated by a dash.

The results in Table 6 should be interpreted as contextual rather than direct head-to-head comparisons because preprocessing procedures, FOV definitions, reference annotations, dataset partitions, and parameter-selection protocols differ across studies. Consequently, statistical superiority cannot be inferred from these comparisons alone. Several supervised or parameter-optimized methods reported higher Dice scores on the shared datasets. For example, Zhou et al. [31] achieved higher overlap on DRIVE and STARE, while

Mahapatra et al. [32] reported higher Dice on HRF. Conversely, the proposed framework exceeded the Dice reported by Mahapatra et al. [32] on STARE. These differences are expected because the compared methods operate under substantially different assumptions. In contrast to supervised or optimization-based approaches, the proposed framework does not require model training, dataset-specific optimization, or learned feature representations.

Table 6. Contextual comparison with representative retinal vessel segmentation methods.

Dataset	Method	Dice	Sensitivity	Specificity	Accuracy
DRIVE	Proposed method	0.7382	0.6705	0.9793	0.9390
	Zhou et al. [31]	0.7781	0.7262	0.9803	0.9475
STARE	Proposed method	0.7286	0.7580	0.9630	0.9422
	Zhou et al. [31]	0.7764	0.7865	0.9730	0.9535
	Mahapatra et al. [32]	0.7129	0.6846	0.9802	0.9601
	Khan et al. [33]	–	0.8078	0.9721	0.9586
HRF	Proposed method	0.6763	0.7847	0.9470	0.9321
	Mahapatra et al. [32]	0.7445	0.7132	0.9814	0.9615

The observed performance therefore reflects the trade-off between interpretability and segmentation accuracy. Fixed local thresholding rules and morphology-based refinement cannot exploit contextual image features as effectively as learned representations, particularly under severe pathological conditions or substantial acquisition variability. Moreover, elongation filtering intentionally suppresses compact false-positive structures to improve precision, but this may remove short vessel fragments, leading to reduced sensitivity and increased HD95. Lesions and abnormal illumination may also satisfy the same local intensity characteristics as retinal vessels, further limiting deterministic rule-based discrimination.

These results indicate that adaptive decision-level late fusion provides competitive and reproducible segmentation performance while maintaining complete interpretability and independence from statistical model fitting. Rather than competing directly with state-of-the-art supervised models under identical optimization settings, the proposed framework offers an alternative operating paradigm for applications where transparency, reproducibility, and freedom from annotated training data are primary design objectives.

4.5. Pathology-Stratified HRF Analysis

To further examine the effect of retinal pathology on segmentation performance, the HRF dataset was analyzed separately for the healthy, glaucomatous, and diabetic retinopathy subgroups. The subgroup-specific results are summarized in Table 7.

Table 7. Pathology-stratified performance of the proposed framework on the HRF dataset.

HRF subgroup (n = 15 each)	Dice	Precision	Sensitivity	HD95 (px)	clDice
Healthy	0.7414	0.6941	0.7972	42.45	0.7818
Glaucoma	0.6618	0.5773	0.7784	53.11	0.7022
Diabetic retinopathy	0.6257	0.5295	0.7786	47.03	0.6833

As shown in Table 7, segmentation performance declined progressively from healthy images to glaucomatous and diabetic retinopathy cases. Dice decreased from 0.7414 for healthy images to 0.6618 for glaucoma and 0.6257 for diabetic retinopathy. The reduction was accompanied primarily by a decrease in precision, whereas sensitivity remained comparatively stable across the three subgroups. HD95 increased for glaucomatous images, while clDice followed the same decreasing trend as Dice, indicating reduced preservation of vascular connectivity under pathological conditions.

The relatively stable sensitivity suggests that the proposed framework continues to identify most major vascular structures even in diseased retinas. Instead, the principal source of performance degradation is the increased number of false-positive responses, reflected by the substantial reduction in precision. Bright lesions, exudates, hemorrhages, and abnormal

illumination patterns may satisfy local intensity and morphological criteria that resemble retinal vessels, causing non-vascular structures to survive the deterministic refinement stages.

These findings identify an important limitation of the proposed rule-based framework. Although the fixed decision rules generalize reasonably well across heterogeneous datasets, they remain sensitive to lesion-induced appearance changes that cannot be distinguished using local intensity and geometric information alone. Consequently, the present results should not be interpreted as evidence of pathology robustness but rather as an indication of the operating conditions under which deterministic segmentation remains effective.

4.6. Qualitative Interpretation and Limitations

Representative qualitative results are presented in Figure 2, including the original retinal fundus image, the CLAHE-enhanced image, the aligned expert annotation, the final segmentation produced by the proposed framework, and the six constituent thresholding outputs. This comparison enables visual assessment of the contribution of adaptive late fusion and morphological refinement beyond the individual thresholding methods.

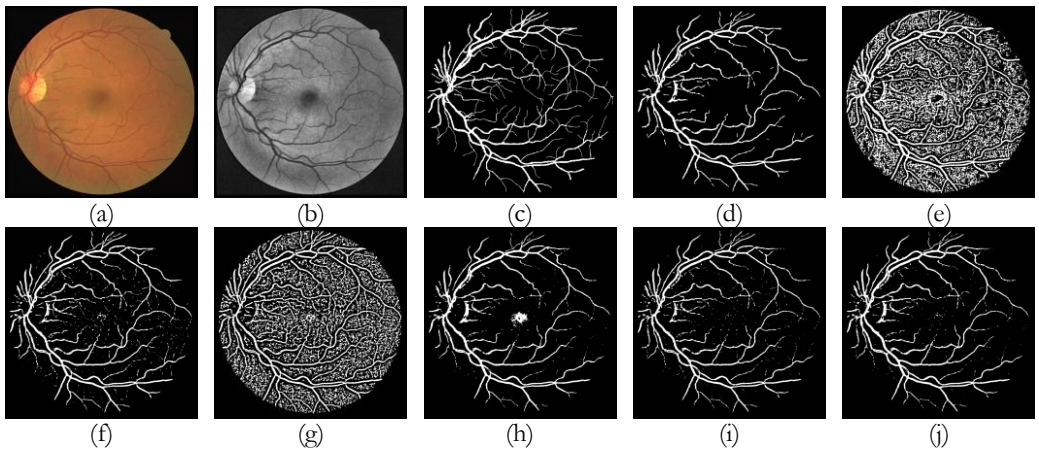

Figure 2. Representative qualitative comparison of retinal vessel segmentation. (**a**) Original retinal fundus image; (**b**) CLAHE-enhanced image; (**c**) aligned reference annotation; (**d**) proposed segmentation; (**e**) Bernsen; (**f**) Bradley; (**g**) Niblack; (**h**) Phansalkar; (**i**) Sauvola; (**j**) Wolf.

The qualitative examples are consistent with the quantitative results reported in the preceding sections. The proposed framework generally preserves the continuity of the principal vascular network while suppressing many isolated false-positive structures that remain visible in the individual threshold outputs. The morphology-guided refinement stages contribute to cleaner vessel boundaries and improved structural consistency, whereas residual errors occur predominantly around pathological regions, weak capillaries, and areas with pronounced illumination variation. These observations are consistent with the improvements in Dice, precision, and clDice obtained after elongation filtering and boundary smoothing, as well as the remaining reductions in sensitivity and boundary accuracy observed during the ablation analysis.

Within its intended operating scope, the proposed framework demonstrates several practical advantages. A single fixed parameter configuration was evaluated across five independent datasets without dataset-specific optimization, the complete framework outperformed every selected constituent thresholding method in pooled Dice, and it improved upon fixed majority voting for 128 of 135 evaluation images. In addition, the explicit intermediate representations—including individual threshold masks, weighted voting maps, reconstruction stages, and rejected connected components—provide complete transparency of the segmentation process, facilitating interpretation, debugging, and reproducibility. These characteristics distinguish the proposed framework from learning-based methods whose internal feature representations are generally less interpretable.

Despite these strengths, several limitations remain. Parameter configuration and fusion weights were determined from only eight development images in the DRIVE training partition, and the weighting strategy remains heuristic. Spatial normalization to a fixed 2,048-pixel retinal diameter may not fully accommodate images with substantially different anatomical

scales. The framework produces binary segmentation masks rather than calibrated confidence maps and exhibits reduced performance on pathological HRF images. Furthermore, the Traced variants depend on an intensity-based optic-disc candidate, and the study does not include prospective clinical validation, common-protocol reimplementation of published methods, or controlled runtime comparisons with optimized learning-based approaches. The parallel thresholding branches and sequential morphological refinement also increase computational cost despite image-level parallel execution.

Future work will investigate nested cross-dataset parameter selection, scale-adaptive spatial parameters, automatic optimization of fusion weights without introducing test-set leakage, more robust anatomical anchoring, leave-one-threshold-out fusion analysis, continuous confidence-map generation, topology-aware refinement, broader evaluation across imaging devices and pathological conditions, hybrid integration with self-supervised learning, computational optimization, and downstream retinal vascular measurements.

Overall, this study demonstrates that adaptive decision-level late fusion can provide a transparent and reproducible alternative to learning-based retinal vessel segmentation when annotated training data are unavailable, or interpretability is a primary design requirement. While the proposed deterministic framework does not aim to replace state-of-the-art supervised models, it establishes a strong, interpretable baseline that can serve as a complementary foundation for future hybrid retinal vessel segmentation methods.

5. Conclusions

This study presented a rule-based unsupervised retinal vessel segmentation framework based on adaptive decision-level late fusion and sequential morphological refinement. The proposed framework combines six complementary local adaptive thresholding methods without statistical model fitting or annotated training data, providing a transparent and fully reproducible segmentation pipeline. Using a strictly held-out evaluation comprising 135 retinal fundus images from five public datasets, the complete framework achieved an image-weighted Dice score of 0.7104 and clDice of 0.7395. The proposed framework consistently outperformed every selected constituent thresholding output and improved Dice for 128 of 135 evaluation images relative to fixed majority voting. The ablation study further demonstrated that the principal performance gain originated from morphology-guided refinement, particularly elongation filtering, while adaptive decision-level weighting alone provided limited benefit. These findings indicate that combining complementary thresholding methods with structural refinement produces more reliable vessel segmentation than any individual thresholding approach.

The results also highlight the practical strengths of the proposed framework. A single fixed parameter configuration generalized across five independent datasets acquired under different imaging conditions without dataset-specific optimization, while maintaining complete interpretability throughout every stage of the segmentation process. Consequently, the proposed framework provides a reproducible classical alternative for retinal vessel segmentation in applications where methodological transparency, deterministic behavior, and independence from trained segmentation models are primary design requirements. Nevertheless, several limitations should be acknowledged. The framework remains dependent on development-set parameter selection, heuristic fusion weights, fixed spatial normalization, and intensity-based optic-disc localization for the Traced variants. Performance also decreases in pathological retinal images, reflecting the inherent limitations of deterministic local intensity rules when confronted with lesion-induced appearance changes. Furthermore, the framework produces binary segmentation masks rather than calibrated confidence estimates and introduces additional computational cost through parallel thresholding and sequential refinement.

Overall, the proposed framework should be regarded as an interpretable and reproducible rule-based segmentation approach rather than a replacement for state-of-the-art learning-based methods or evidence of clinical readiness. The future research directions discussed in Section 4.6 provide a roadmap for extending the present framework toward more adaptive, topology-aware, and hybrid segmentation strategies.

Author Contributions: Conceptualization: A.A.; Methodology: A.A. and M.B.A.M.; Software: A.A.; Validation: M.B.A.M., A.A.N., and H.R.E.; Formal analysis: A.A. and A.A.N.;

Investigation: A.A. and H.R.E.; Resources: A.A.; Data curation: A.A.; Writing—original draft preparation: A.A.; Writing—review and editing: M.B.A.M., A.A.N., and H.R.E.; Visualization: A.A.; Supervision: M.B.A.M.; Project administration: A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universitas Amikom Yogyakarta

Data Availability Statement: The datasets used in this study are publicly available in https://github.com/girfa/RetinalVesselDatasets. The MATLAB implementation, fixed parameter configuration, and supporting evaluation results are available from the corresponding author upon reasonable request.

Acknowledgments: The authors acknowledge the use of AI-assisted technologies (Large Language Models) for language refinement and structural editing during the preparation of this manuscript. The authors have thoroughly reviewed and edited all generated text and assume full responsibility for the final published content.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, the collection, analysis, or interpretation of data, the writing of the manuscript, or the decision to publish the results.

References

[1] I. Inamullah, I. Razzak, and S. Jameel, "The eye as a window to systemic health: A survey of retinal imaging from classical techniques to oculomics," *J. Precis. Med. Heal. Dis.*, vol. 4, p. 100023, Dec. 2025, doi: 10.1016/j.premed.2025.100023.

[2] M. N. Getahun, O. Y. Rogov, D. V. Dylov, A. Somov, A. Bouridane, and R. Hamoudi, "FS-Net: Full scale network and adaptive threshold for improving extraction of micro-retinal vessel structures," *Pattern Recognit. Lett.*, vol. 189, pp. 188–194, Mar. 2025, doi: 10.1016/j.patrec.2025.01.019.

[3] N. Popovic, M. Radunovic, J. Badnjar, and T. Popovic, "Fractal dimension and lacunarity analysis of retinal microvascular morphology in hypertension and diabetes," *Microvasc. Res.*, vol. 118, pp. 36–43, Jul. 2018, doi: 10.1016/j.mvr.2018.02.006.

[4] F. Yan, Y. Liu, Q. Zhao, and G. He, "A multi-task deep learning pipeline integrating vessel segmentation and radiomics for multiclass retinal disease classification," *Photodiagnosis Photodyn. Ther.*, vol. 56, p. 105209, Dec. 2025, doi: 10.1016/j.pdpdt.2025.105209.

[5] Z. Ma and X. Li, "An improved supervised and attention mechanism-based U-Net algorithm for retinal vessel segmentation," *Comput. Biol. Med.*, vol. 168, p. 107770, Jan. 2024, doi: 10.1016/j.compbiomed.2023.107770.

[6] Y. Wu, C. S. Lee, and A. Y. Lee, "Beyond convolutions and supervised learning with transformers and representation learning for retinal image analysis," *Prog. Retin. Eye Res.*, vol. 110, p. 101419, Jan. 2026, doi: 10.1016/j.preteyeres.2025.101419.

[7] J. Liu *et al.*, "Unsupervised domain adaptation multi-level adversarial learning-based crossing-domain retinal vessel segmentation," *Comput. Biol. Med.*, vol. 178, p. 108759, Aug. 2024, doi: 10.1016/j.compbiomed.2024.108759.

[8] W. Ben Messaoud, S. Ben Brahim, A. Zrelli, and M. Yahia, "Domain-specific data augmentation for retinal disease classification using hybrid CNN-Transformer architectures," *Biomed. Signal Process. Control*, vol. 121, p. 110325, Aug. 2026, doi: 10.1016/j.bspc.2026.110325.

[9] J. Van Eijgen *et al.*, "Retinal vessel analysis to assess microvascular function in the healthy eye: A systematic review on the response to acute physiological and pathological stressors," *Surv. Ophthalmol.*, vol. 70, no. 2, pp. 200–214, Mar. 2025, doi: 10.1016/j.survophthal.2024.11.008.

[10] Z. Li, H. Wang, N. Tang, R. Lan, B. Li, and X. Luo, "A Dual-Layer Semantic Fusion Network for Retinal Vessel Segmentation," *Biomed. Signal Process. Control*, vol. 112, p. 108463, Feb. 2026, doi: 10.1016/j.bspc.2025.108463.

[11] N. Alsharabi, A. Echtioui, A. Alayba, G. Alshammari, and H. Hamam, "Advanced hybrid UNet architectures for retinal vessel segmentation," *Egypt. Informatics J.*, vol. 32, p. 100823, Dec. 2025, doi: 10.1016/j.eij.2025.100823.

[12] H. Hu, H. Xiong, F. Long, M. S. Alam, and J. Sang, "Multi-scale Vision Mamba-UNet: A Mamba-based method for retinal vessel segmentation," *Biomed. Signal Process. Control*, vol. 112, p. 108435, Feb. 2026, doi: 10.1016/j.bspc.2025.108435.

[13] S. Aslani and H. Sarnel, "A new supervised retinal vessel segmentation method based on robust hybrid features," *Biomed. Signal Process. Control*, vol. 30, pp. 1–12, Sep. 2016, doi: 10.1016/j.bspc.2016.05.006.

[14] S. Das, S. Chakraborty, M. Mishra, and S. Majumder, "Assessment of retinal blood vessel segmentation using U-Net model: A deep learning approach," *Franklin Open*, vol. 8, p. 100143, Sep. 2024, doi: 10.1016/j.fraope.2024.100143.

[15] L. Ma, J.-X. Jiang, S. Zeng, Y. Li, and Z. Wang, "Frequency view feature interpolation and disentanglement for single-source domain generalization in retinal vessel segmentation," *Biomed. Signal Process. Control*, vol. 112, p. 108814, Feb. 2026, doi: 10.1016/j.bspc.2025.108814.

[16] S. Guo, "Enhancing cross-domain generalization in retinal image segmentation via style randomization and style normalization," *Image Vis. Comput.*, vol. 162, p. 105694, Oct. 2025, doi: 10.1016/j.imavis.2025.105694.

[17] S. Guo and K. Ji, "Random color transformation for single domain generalized retinal image segmentation," *Eng. Appl. Artif. Intell.*, vol. 136, p. 108907, Oct. 2024, doi: 10.1016/j.engappai.2024.108907.

[18] D. Relan and R. Relan, "Unsupervised sorting of retinal vessels using locally consistent Gaussian mixtures," *Comput. Methods Programs Biomed.*, vol. 199, p. 105894, Feb. 2021, doi: 10.1016/j.cmpb.2020.105894.

[19] B. Subramani, M. Veluchamy, and A. K. Bhandari, "Optimal color constancy model for biological feature enhancement of retinal fundus images," *Biomed. Signal Process. Control*, vol. 117, p. 109637, May 2026, doi: 10.1016/j.bspc.2026.109637.
[20] Q. Xin and L. Han, "Dynamic spatial graph convolution and attention network with dual threshold inference for retinal vessel segmentation," *Biomed. Signal Process. Control*, vol. 120, p. 110158, Jul. 2026, doi: 10.1016/j.bspc.2026.110158.
[21] M. Nandy and M. Banerjee, "A Comparative Analysis of Application of Niblack and Sauvola Binarization to Retinal Vessel Segmentation," in *2017 3rd International Conference on Computational Intelligence and Networks (CINE)*, Oct. 2017, pp. 105–109. doi: 10.1109/CINE.2017.19.
[22] R. K. Hong *et al.*, "Quantifying the repeatability and reproducibility of Dr. Noon CVD, AI software as medical device for cardiovascular risk assessment via retinal imaging," *Can. J. Ophthalmol.*, vol. 61, no. 1, pp. 1–11, Feb. 2026, doi: 10.1016/j.jcjo.2025.07.008.
[23] B. C. Fry *et al.*, "Predicting the impact of retinal vessel density on retinal vessel and tissue oxygenation using a theoretical model," *Math. Biosci.*, vol. 377, p. 109292, Nov. 2024, doi: 10.1016/j.mbs.2024.109292.
[24] X. Shi *et al.*, "The retinal nerve fiber layer mean thickness in patients with early Parkinson's disease reflects striatal dopamine function," *Neuroscience*, vol. 599, pp. 94–102, Apr. 2026, doi: 10.1016/j.neuroscience.2026.02.013.
[25] Z. Chu *et al.*, "Quantification of Choriocapillaris with Phansalkar Local Thresholding: Pitfalls to Avoid," *Am. J. Ophthalmol.*, vol. 213, pp. 161–176, May 2020, doi: 10.1016/j.ajo.2020.02.003.
[26] D. Benny and K. R. Soumya, "New local adaptive thresholding and dynamic self-organizing feature map techniques for handwritten character recognizer," in *2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2015]*, Mar. 2015, pp. 1–4. doi: 10.1109/ICCPCT.2015.7159293.
[27] T. Mapayi, S. Viriri, and J.-R. Tapamo, "A New Adaptive Thresholding Technique for Retinal Vessel Segmentation Based on Local Homogeneity Information," in *Lecture Notes in Computer Science*, 2014, pp. 558–567. doi: 10.1007/978-3-319-07998-1_64.
[28] L. Du *et al.*, "Deep ensemble learning for accurate retinal vessel segmentation," *Comput. Biol. Med.*, vol. 158, p. 106829, May 2023, doi: 10.1016/j.compbiomed.2023.106829.
[29] M. Zhang, J. Wang, X. Cao, X. Xu, J. Zhou, and H. Chen, "An integrated global and local thresholding method for segmenting blood vessels in angiography," *Heliyon*, vol. 10, no. 22, p. e38579, Nov. 2024, doi: 10.1016/j.heliyon.2024.e38579.
[30] Xiaoyi Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 1, pp. 131–137, Jan. 2003, doi: 10.1109/TPAMI.2003.1159954.
[31] C. Zhou, X. Zhang, and H. Chen, "A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden Markov model," *Comput. Methods Programs Biomed.*, vol. 187, p. 105231, Apr. 2020, doi: 10.1016/j.cmpb.2019.105231.
[32] S. Mahapatra, S. Agrawal, P. K. Mishro, and R. B. Pachori, "A novel framework for retinal vessel segmentation using optimal improved frangi filter and adaptive weighted spatial FCM," *Comput. Biol. Med.*, vol. 147, p. 105770, Aug. 2022, doi: 10.1016/j.compbiomed.2022.105770.
[33] T. M. Khan *et al.*, "Width-wise vessel bifurcation for improved retinal vessel segmentation," *Biomed. Signal Process. Control*, vol. 71, p. 103169, Jan. 2022, doi: 10.1016/j.bspc.2021.103169.
[34] Z. Qu, L. Zhuo, J. Cao, X. Li, H. Yin, and Z. Wang, "TP-Net: Two-Path Network for Retinal Vessel Segmentation," *IEEE J. Biomed. Heal. Informatics*, vol. 27, no. 4, pp. 1979–1990, Apr. 2023, doi: 10.1109/JBHI.2023.3237704.
[35] D. Hu, H. Li, H. Liu, and I. Oguz, "Domain generalization for retinal vessel segmentation via Hessian-based vector field," *Med. Image Anal.*, vol. 95, p. 103164, Jul. 2024, doi: 10.1016/j.media.2024.103164.
[36] Z. Tu, Q. Zhou, H. Zou, and X. Zhang, "A Multi-Task Dense Network with Self-Supervised Learning for Retinal Vessel Segmentation," *Electronics*, vol. 11, no. 21, p. 3538, Oct. 2022, doi: 10.3390/electronics11213538.
[37] L. Tong *et al.*, "LiViT-Net: A U-Net-like, lightweight Transformer network for retinal vessel segmentation," *Comput. Struct. Biotechnol. J.*, vol. 24, pp. 213–224, Dec. 2024, doi: 10.1016/j.csbj.2024.03.003.
[38] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-Based Vessel Segmentation in Color Images of the Retina," *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 501–509, Apr. 2004, doi: 10.1109/TMI.2004.825627.
[39] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imaging*, vol. 19, no. 3, pp. 203–210, Mar. 2000, doi: 10.1109/42.845178.
[40] M. M. Fraz *et al.*, "An Ensemble Classification-Based Approach Applied to Retinal Blood Vessel Segmentation," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2538–2548, Sep. 2012, doi: 10.1109/TBME.2012.2205687.
[41] J. Odstrcilik *et al.*, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database," *IET Image Process.*, vol. 7, no. 4, pp. 373–383, Jun. 2013, doi: 10.1049/iet-ipr.2012.0455.
[42] J. I. Orlando, J. Barbosa Breda, K. van Keer, M. B. Blaschko, P. J. Blanco, and C. A. Bulant, "Towards a Glaucoma Risk Index Based on Simulated Hemodynamics from Fundus Images," in *Lecture Notes in Computer Science*, 2018, pp. 65–73. doi: 10.1007/978-3-030-00934-2_8.
[43] S. Shit *et al.*, "clDice - a Novel Topology-Preserving Loss Function for Tubular Structure Segmentation," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2021, pp. 16555–16564. doi: 10.1109/CVPR46437.2021.01629.