International Journal of Applied Science; Vol. 8, No. 3; 2025
ISSN 2576-7240 E-ISSN 2576-7259
https://doi.org/10.30560/ijas.v8n3p98

Retinal Vessel Segmentation Based on Feature Fusion and Spatially Constrained Co-Clustering

Chukun Chen[1]

[1] Jieyang Polytechnic, China

Correspondence: Chen Chukun, Jieyang Polytechnic, Jieyang City, Guangdong Province, China.

Received: September 15, 2025; Accepted: September 26, 2025; Published: September 27, 2025

This paper presents the research results supported by the 2023 Scientific Research Project of Jieyang Polytechnic (2023JYCKY06) and the 2023-2025 Innovation and Strengthening Higher Education Institutions Project of Jieyang Polytechnic (JYPCQKY23-0518).

Abstract

Retinal vessel segmentation plays a critical role in the early diagnosis of ocular and systemic diseases such as diabetic retinopathy and hypertension. This paper proposes an unsupervised segmentation method that integrates dual-feature fusion and spatially constrained co-clustering. The approach combines multi-scale and multi-directional vascular structural information captured by B-COSFIRE filters with local contrast enhancement from the Top-hat transformation, resulting in a complementary feature representation. Furthermore, a spatial constraint based on Markov Random Fields is incorporated into an information-theoretic co-clustering framework to improve topological consistency. Experimental results on the STARE database demonstrate that the proposed method achieves competitive performance, with an accuracy of 0.9511, sensitivity of 0.7439, and specificity of 0.9747, effectively enhancing vascular continuity and robustness.

Keywords: retinal vessel segmentation, Co-clustering, B-COSFIRE filters, top-hat transformation, spatial constraints, Markov Random Field

1. Introduction

Morphological changes in retinal vessels serve as a key early diagnostic basis for various vascular diseases, such as diabetic retinopathy and hypertensive retinopathy, as indicated by Liu et al. (2024)[1]. With the continuous rise in the global prevalence of diabetes, the demand for automated screening of diabetic retinopathy has become increasingly urgent. Joseph et al. (2024)[2] emphasized that accurate retinal vessel segmentation is a crucial foundation for realizing the automated diagnosis of such lesions.

In recent years, retinal vessel segmentation technology has presented a trend of parallel development between traditional image processing methods and deep learning methods. Qin, Q., & Chen, Y. (2024)[3], in their review study, pointed out that deep learning methods have become a research focus due to their end-to-end feature extraction capability. For example, Yin et al. (2021)[4] proposed an end-to-end neural network model with a nested structure (MS-UNet++), which introduces a MultiRes module to enhance feature learning for tiny vessels in low-contrast environments, and combines a SENet module to perform channel-wise weighting on features—this helps expand the receptive field and strengthen target-related features. Almarru et al. (2024)[5] developed GANVesselNet, which adopts a dual-path Generative Adversarial Network (GAN) structure, significantly improving the detection sensitivity for tiny vessels. However, deep learning methods still have several limitations. First, Ouyang et al. (2023)[6] noted that complex models often have a high deployment threshold. Their proposed LEA U-Net enhances the recognition performance of tiny vessels through a local feature enhancement module, but its large parameter scale and slow inference speed make it difficult to apply to resource-constrained clinical terminal devices. Second, Liu et al. (2024)[1] also argued that such methods have insufficient adaptability to complex pathological features in clinical practice, such as vascular tortuosity and stenosis. Moreover, Yin et al. (2021)[4]'s research showed that improvements in model performance are often accompanied by a sharp increase in computational complexity, leading to difficulties in balancing accuracy and efficiency. Additionally, deep models rely on high-quality data annotated by professional physicians for supervised learning.

Compared with deep learning methods, Wang and Zhang (2019)[3] stressed that traditional image processing algorithms still have irreplaceable advantages in application scenarios with high real-time requirements, as they feature high computational efficiency and do not depend on annotated data. Among traditional methods, Chaudhuri et al. (1989)[7] proposed matched filtering as one of the early representative technologies—they used a two-dimensional Gaussian template to simulate the vascular cross-sectional structure and enhance the response of vascular regions. However, this method only relies on a single grayscale feature, making it unable to adapt to the multi-directional and multi-scale changes of blood vessels. Morphological methods, another important branch of traditional techniques, aim to enhance vascular morphology through structural element operations. Nandy et al. (2023)[8] proposed the Bel-Hat transformation, which combines neighborhood-adaptive line structural elements and two-dimensional Gaussian structural elements to effectively improve the contrast between vessels and the background. Nevertheless, its threshold-based segmentation strategy is prone to causing vascular discontinuities when processing low-contrast regions.

In existing research, Wang and Zhang (2019)[3] summarized that traditional methods generally face two bottlenecks. First, their limited feature expression capability makes it difficult to fully capture the diversity of blood vessels in terms of grayscale distribution and morphological orientation; second, the lack of modeling for spatial contextual relationships leads to pixel-independent decision-making methods being easily affected by noise interference, resulting in noise in segmentation results. To improve the continuity of vascular structures, Azzopardi et al. (2015)[9] proposed the B-COSFIRE filter, which has a specific response to rod-like structures and can effectively detect main blood vessels, but its sensitivity to tiny vessels is still insufficient. Zheng et al. (2025)[10] combined B-COSFIRE with phase congruency features and introduced clustering methods for segmentation, which improved performance to a certain extent. However, they still failed to completely solve the problems of multi-feature complementarity and spatial contextual collaborative modeling, making it hard to achieve an ideal balance between segmentation accuracy and vascular continuity.

To overcome the aforementioned limitations of traditional methods and deep learning methods, this paper proposes a retinal vessel segmentation algorithm that integrates dual-feature extraction and spatially constrained co-clustering. The innovative value of this method is mainly reflected in the following two aspects: First, by fusing the advantages of Top-hat transformation in enhancing local brightness differences and the capability of B-COSFIRE filters in multi-directional vascular structure perception, a complementary dual-feature system is constructed to more comprehensively characterize the grayscale distribution and morphological orientation characteristics of blood vessels. Second, based on the information-theoretic co-clustering framework proposed by Dhillon et al. (2003)[12], a spatial constraint mechanism is introduced and combined with the Markov Random Field (MRF) model proposed by Geman et al. (1984)[13]. By fusing the neighborhood spatial information of pixels, this mechanism suppresses noise interference while improving the connectivity and structural continuity of the vascular network. Experimental results show that the proposed method can significantly enhance the detection performance for blood vessels with different orientations and low-contrast tiny vessels, ultimately obtaining retinal vessel segmentation results with higher structural integrity.

2. Algorithm

2.1 Preprocessing

To reduce noise interference in raw retinal images and enhance vascular structures, this study first performs preprocessing operations. First, the green channel of the image is extracted. Since the grayscale contrast between blood vessels and background tissues is most significant in this channel, this method can effectively suppress the masking of vascular information by noise in other channels. Subsequently, Contrast-Limited Adaptive Histogram Equalization (CLAHE) is applied to enhance the green channel image. While suppressing the amplification of local noise, this technique enhances the grayscale details of tiny capillaries, thereby improving the distinguishability between blood vessels and the background.

2.2 Feature Extraction

2.2.1 B-COSFIRE Filter

The B-COSFIRE filter (bar-selective combination of shifted filter responses) is a rod-shaped structure detection tool based on the response mechanism of visual cortex cells, which is suitable for the detection of blood vessel tubular structures in fundus images. Its core principle is to achieve precise enhancement of blood vessels of different shapes by combining multi - scale Difference of Gaussian (DoG) filter responses. The filter first detects the key points with significant intensity changes in the image through the DoG operator, and these key points correspond to the characteristic positions of the blood vessel edges. For each key point, fuzzy and shift operations are used to process the filter responses to enhance fault - tolerance, and then multiple groups of filter responses

are aggregated through weighted geometric mean. The specific response formulas are shown in Formulas (1) and (2).

$$B(x,y)=\left|\left\{\prod_{i=1}^{V}\left[V_{\sigma_i,\rho_i,\phi_i}(x,y)\right]^{\omega_i}\right\}^{1/\sum_{i=1}^{|V|}\omega_i}\right|_t \tag{1}$$

$$\omega_i=\exp\left(-\frac{\rho_i^2}{2\hat{\sigma}^2}\right),\hat{\sigma}=\frac{1}{3}\max_{i\in\{1,\ldots,|V|\}}\{\rho_i\} \tag{2}$$

Among them, $V_{\sigma_i,\rho_i,\varphi_i}(x,y)$ represents the response value of the i-th filter; σ_i denotes the standard deviation of the Gaussian function corresponding to the maximum filter response; ρ_i and φ_i are the polar coordinates relative to the support center of the filter. The set $V=\{(\sigma_i,\rho_i,\varphi_i)\mid i=1,\ldots,n\}$ records the information of each key point, where g stands for the number of filters.

To adapt to the multi-directional distribution characteristics of retinal blood vessels, this study rotates the filter in 12 equally spaced directions, and finally takes the maximum value of the responses across all directions as the output feature.

2.2.2 Top-hat Transformation

As a classic morphological brightness enhancement tool, the Top-Hat transformation effectively extracts small structures in an image with grayscale values higher than the background by calculating the difference between the original image and the result of the opening operation. It is particularly suitable for addressing the issue where low-contrast small blood vessels in retinal images are easily obscured by the background, while also exhibiting strong robustness against slowly varying uneven illumination. This method is based on the mathematical morphology theory proposed by Jean Serra, and its core formula is defined as Formula (3).

$$TopHat(A)=A-(A\circ B) \tag{3}$$

where A denotes the input image, B is the structuring element, and $\circ$ represents the opening operation. The opening operation itself is defined as a composite operation of erosion followed by dilation: $(A\circ B)=(A\Theta B)\oplus B$ (where Θ stands for the erosion operation and $\oplus$ denotes the dilation operation).

In terms of structuring element selection, this study adopts a disk-shaped structuring element (se_r, where r denotes the radius). Its circular contour can better match the tubular cross-sectional characteristics of blood vessels, effectively avoiding the problem of insufficient enhancement for oblique blood vessels caused by directional deviation.

In terms of the multi-scale strategy, a series of structuring elements with radii $R=[1,3]$ are selected to perform the opening operation. Small radii are used to capture the details of capillaries, while large radii are employed to retain the structure of main blood vessels, thus achieving full coverage of blood vessels with different diameters.

Finally, a feature map is generated through a response aggregation mechanism, and the Top-hat feature value of each pixel is defined as:

$$T(x,y)=\max_{n\in^R}[I(x,y)-(I(x,y)\circ se_r)] \tag{4}$$

where $R=[1,3]$ is the radius set, and $I(x,y)$ represents the grayscale value of the preprocessed image. This feature forms a precise complement to the B-COSFIRE feature. The former focuses on the local brightness differences between blood vessels and the background, while the latter excels at capturing the directional and scale characteristics of blood vessels. The combination of the two can comprehensively characterize the morphological features of retinal blood vessels.

2.3 Spatially Constrained Co-Clustering

This section elaborates on the core algorithm of this study. The spatially constrained information-theoretic co-clustering model. Innovatively, this model achieves the co-clustering of pixels and dual features (B-COSFIRE, Top-hat) in a joint manner, enabling in-depth exploration of the intrinsic correlation patterns between features and

pixel categories. Meanwhile, it incorporates the spatial constraint mechanism of MRF. This effectively addresses the issues of insufficient topological continuity and noise sensitivity that traditional clustering methods face in retinal vessel segmentation, ultimately realizing accurate and robust extraction of vascular structures.

2.3.1 Problem Formulation and Input Definition

The preprocessed retinal image is constrained to the region defined by its Field of View (FOV) mask, containing N valid pixels requiring classification. For each pixel (x, y), two key features are extracted. The B-COSFIRE orientation response feature $B(x, y)$, which characterizes the saliency of tubular structures, and the Top-hat transformation intensity feature $T(x, y)$, designed to enhance bright linear structures. These features are organized into a feature matrix $X \in R^{N\times 2}$, serving as the input to the subsequent clustering algorithm.

This study formulates the vessel segmentation task as a co-clustering problem, aiming to simultaneously achieve dual partitioning of both pixels and features:

(1) Pixel Clustering. Assign the N pixels into $K = 2$ disjoint clusters, denoted as $c_p = \{c_{p1}, c_{p2}\}$, where c_{p1} represents the vessel pixel class and c_{p2} corresponds to the background pixel class. This process directly achieves binary segmentation of vessels versus background.

(2) Feature Clustering. Partition the two features into $L = 2$ functional clusters, denoted as $c_f = \{c_{f1}, c_{f2}\}$, intended to identify distinct functional categories of features (e.g., orientation feature group and intensity feature group) and reveal inherent feature grouping patterns.

The principal advantage of this joint partitioning strategy is its capability to discover statistical dependencies between feature combinations and pixel categories. For instance, the algorithm can automatically identify that a feature combination pattern exhibiting strong B-COSFIRE response coupled with strong Top-hat response is highly associated with the vessel pixel category. This mechanism provides more robust and interpretable decision-making criteria for pixel classification, thereby significantly enhancing the accuracy of vessel segmentation.

2.3.2 Fundamentals of Information-Theoretic Co-Clustering

Information-Theoretic Co-Clustering(ITCC) serves as the core theoretical framework, whose essence lies in achieving optimal partitioning by maximizing the mutual information(MI) between pixel clusters c_p and feature clusters c_f. Mutual information, which quantifies the degree of dependency between two random variables, is mathematically defined by Equation (5).

$$I(c_p; c_f) = \sum_{k=1}^{K}\sum_{l=1}^{L} P(c_{pk}, c_{fl}) \log\left(\frac{P(c_{pk}, c_{fl})}{P(c_{pk})P(c_{fl})}\right) \tag{5}$$

where $P(c_{pk}, c_{fl})$ denotes the joint probability distribution of pixel cluster c_{pk} and feature cluster c_{fl}, while $P(c_{pk})$ and $P(c_{fl})$ represent their marginal probability distributions, respectively. Maximizing mutual information implies that, when the feature cluster categories are known, the pixel cluster categories can be inferred with maximum accuracy, and vice versa. Thereby, it ensures that the guiding value of features for pixel classification is maximized.

The standard ITCC algorithm employs an alternating iterative optimization strategy to maximize the objective function. First, the feature cluster assignment c_f is fixed, and the pixel cluster assignment c_p is adjusted to

increase the mutual information $I(c_p;c_f)$. Then, the pixel cluster assignment c_p is fixed, and the feature cluster assignment c_f is re-optimized to further enhance the mutual information.

However, the inherent spatial correlation in retinal images, where adjacent pixels are significantly more likely to belong to the same tissue class than would be expected by chance, is not considered in the standard ITCC framework. This omission leads to heightened sensitivity to image noise, making the segmentation results prone to issues such as vessel fragmentation and salt-and-pepper noise.

2.3.3 Incorporation of Spatial Constraints

To address the limitation of standard co-clustering ignoring spatial contextual information, this study introduces a spatial smoothness constraint term $R(c_p)$ based on a MRF. This constraint enhances the spatial continuity of the segmentation result by penalizing label disagreements between adjacent pixels. It is mathematically defined as follows:

$$R(c_p)=\sum_{i=1}^{N}\sum_{j\in N_i}\omega_{ij}\cdot\delta(c_p^i\neq c_p^j) \tag{6}$$

The physical meaning and design rationale of each parameter are as follows:

(1) The neighborhood system N_i adopts an 8-neighborhood structure (including up, down, left, right, and the four diagonal directions). Compared to a 4-neighborhood, this configuration more comprehensively captures the diagonal continuity of vessels, making it particularly suitable for the complex branching topology of retinal vasculature.

(2) The penalty weight ω_{ij} is uniformly set to 1. Experimental validation has shown that in retinal images, neighboring pixels in all directions contribute equally to vascular continuity. Introducing directional weighting could potentially impose overly strong prior biases.

(3) The indicator function $\delta(\cdot)$ returns a value of 1 when adjacent pixels belong to different clusters, thereby applying a unit penalty. This mechanism aligns with the energy minimization principle of MRF, encouraging spatial smoothness by increasing the energy cost associated with label inconsistencies.

The introduction of this constraint term enables the model to leverage the spatial prior knowledge inherent in retinal images, effectively suppressing isolated misclassifications caused by noise. It leads to notable improvements in continuity, particularly within regions of fine vascular structures.

2.3.4 Joint Objective Function and Optimization

The spatial constraint term is integrated with the information-theoretic co-clustering objective function to construct a joint optimization framework that balances feature relevance and spatial continuity:

$$(c_p^*,c_f^*)=\arg\max[I(c_p;c_f)-\lambda\cdot R(c_p)] \tag{7}$$

where the non-negative hyperparameter controls the relative importance between feature mutual information and spatial smoothness. A larger λ places greater emphasis on spatial continuity, making the model suitable for heavily noise-corrupted images. A smaller λ increases reliance on feature discriminability, which is preferable for high-contrast images. In this study, $\lambda=0.9$ was selected, as it achieved the best balance between noise suppression and detail preservation in our experiments.

The joint objective function is optimized using an alternating iterative strategy, with the specific procedure as follows:

(1) Initialization Phase: Randomly assign pixel cluster labels c_p and feature cluster labels c_f, ensuring the

initial assignments cover the basic distribution of both categories.

(2) Iterative Optimization Phase: Repeat the following steps until convergence:

Pixel Cluster Optimization (with c_f fixed): For each pixel i, compute the total cost of assigning it to the vessel cluster or the background cluster:

$$\mathrm{Cos}t(i,c) = -\Delta I_c(c_p;c_f) + \lambda \cdot \sum_{j \in N} \delta(c \neq c_p^j) \tag{8}$$

Here, ΔI_c denotes the incremental gain in mutual information resulting from assigning pixel i to cluster c, computed via local updates of the joint probability distribution. This cost function enforces a dual constraint: the first term maximizes feature discriminability, while the second minimizes spatial inconsistency. Each pixel is ultimately assigned to the cluster with the lowest total cost.

Feature Cluster Optimization (with c_p fixed): For each feature, recompute the mutual information gain when assigned to different feature clusters. The assignment that maximizes $I(c_p;c_f)$ is selected. This step remains consistent with standard ITCC to preserve flexibility in feature grouping.

(3) Convergence and Output Phase: Upon convergence, the final pixel cluster assignment c_p^* is output as the vessel segmentation result. The feature cluster assignment c_f^* is also recorded for subsequent feature importance analysis.

Through this coordinated optimization mechanism, the model simultaneously leverages the complementary information from dual features and the spatial structural information of the image, providing a retinal vessel segmentation solution that achieves both high accuracy and spatial continuity.

2.4 Vessel Identification and Post-processing

Upon completion of the clustering process, the vessel cluster identification is performed by leveraging the specific responsiveness of the B-COSFIRE feature to tubular structures. The average B-COSFIRE response value is computed for each pixel cluster, and the cluster with the highest average response is identified as the vessel cluster, while the remaining clusters are labeled as background, thereby generating an initial binary segmentation result.

To further enhance the segmentation quality, post-processing is applied. First, a morphological closing operation is performed to smooth the vascular regions. This composite operation, consisting of erosion followed by dilation, fills small internal holes within vessels and repairs discontinuities caused by noise. Subsequently, an area filter is applied to eliminate isolated regions with excessively small areas in the segmentation result. After these post-processing refinements, the final segmentation result is obtained.

3. Experimental Results and Discussion

3.1 Experimental Test Database

This study employs the internationally recognized STARE database to validate the performance of the proposed algorithm. Established by the University of California, San Diego, the STARE database is one of the benchmark datasets widely used for evaluating the robustness of retinal vessel segmentation algorithms (Hoover et al., 2000)[14]. It consists of 20 color fundus images with a uniform resolution of 700 × 605 pixels, covering both normal and pathological retinal scenarios. Among these, 10 images exhibit pathological features such as diabetic retinopathy and macular degeneration, while the other 10 belong to normal retinas.

3.2 Evaluation Metrics

To objectively and comprehensively evaluate the performance of retinal vessel segmentation, this study adopts three widely used evaluation metrics in the field of medical image segmentation: Accuracy, Sensitivity, and Specificity.

Accuracy(Acc) measures the overall correctness of pixel classification, defined as the ratio of correctly classified pixels (including both correctly segmented vessel pixels and correctly identified background pixels) to the total number of valid pixels.

Sensitivity(Sen) evaluates the algorithm's ability to detect vessel pixels, calculated as the proportion of correctly segmented vessel pixels relative to all vessel pixels in the ground truth.

Specificity(Spe) assesses the algorithm's ability to suppress background pixels, represented as the ratio of correctly identified background pixels to all background pixels in the ground truth.

The formulas for these metrics are given in Equations (9)–(11):

$$Acc = \frac{TP + TN + FP + FN}{TP + TN} \tag{9}$$

$$Sen = \frac{TP + FN}{TP} \tag{10}$$

$$Spe = \frac{TN + FP}{TN} \tag{11}$$

where, TP (True Positive) denotes the number of correctly segmented vessel pixels; TN (True Negative) denotes the number of correctly identified background pixels; FP (False Positive) denotes the number of background pixels incorrectly classified as vessels; FN (False Negative) denotes the number of vessel pixels missed and classified as background.

3.3 Experimental Results and Analysis

3.3.1 Experimental Results

The proposed algorithm was tested on the STARE database, and the performance was evaluated using the metrics described in Section 3.2. The average (Acc, Sen, and Spe across all 20 images were 0.9511, 0.7439, and 0.9747, respectively, demonstrating strong segmentation performance.

The segmentation result of one representative image from the dataset is illustrated in Figure 1. Subfigure (a) shows the original color fundus image, subfigure (b) presents the manual annotation by an expert, and subfigure (c) displays the segmentation result obtained by the proposed method. It can be observed that the algorithm successfully extracts the major vascular structures consistent with the ground truth and also performs well in detecting fine vessel branches.

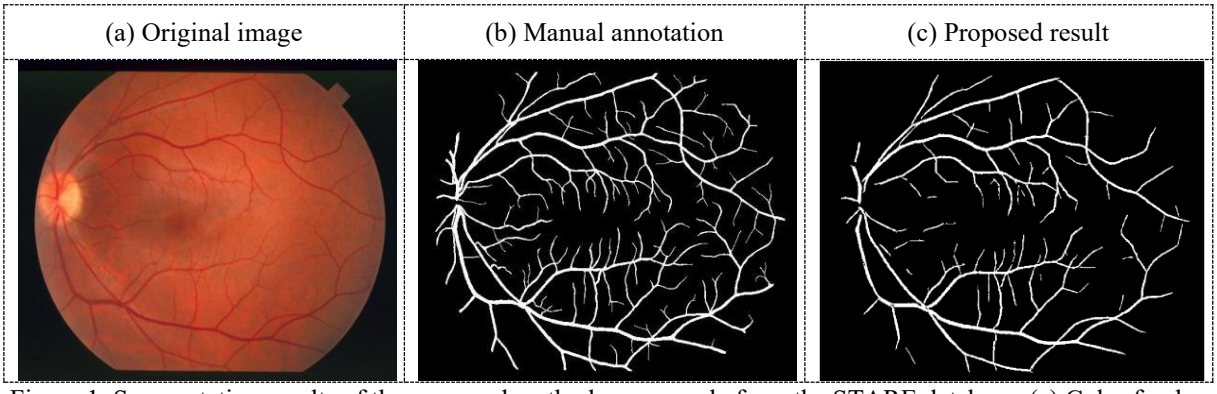

Figure 1. Segmentation results of the proposed method on a sample from the STARE database. (a) Color fundus image; (b) Expert-labeled ground truth; (c) Segmentation result obtained by the proposed algorithm.

3.3.2 Comparative Experimental Analysis

An ablation study was conducted using a controlled variable approach to evaluate the contributions of different feature combinations and the spatial constraint term. The results are summarized in Table 1.

Table 1. Performance comparison of different configurations of the proposed method on the STARE database.

Method	Acc	Se	Sp
B-COSFIRE only	0.9447	0.6508	0.9775
Top-hat only	0.9066	0.8250	0.9164
Dual features without spatial constraint	0.9504	0.7409	0.9744
Proposed method (with spatial constraint)	0.9511	0.7439	0.9747

The analysis of Table 1 reveals distinct performance characteristics across different feature configurations. When using only the B-COSFIRE feature, high specificity is achieved, indicating effective background suppression, but the relatively low sensitivity reflects limited ability to detect vessel pixels. In contrast, the Top-hat feature alone yields improved sensitivity at the cost of significantly reduced specificity, suggesting a higher false positive rate. The combination of both features—even without spatial constraints—results in a more balanced performance, with an accuracy of 0.9504 and competitive sensitivity and specificity, demonstrating the complementary roles of structural and intensity-based features in vessel segmentation. The introduction of the spatial smoothness constraint further enhances all three metrics, albeit marginally, with the most noticeable sensitivity improvement, indicating better recovery of true vessel pixels in structured or noisy regions. The stable specificity confirms that the spatial term does not introduce additional false positives. Overall, these results validate the effectiveness of the proposed dual-feature design with spatial regularization in achieving a superior balance between vessel detection and background suppression compared to single-feature baselines.

The visual comparison of segmentation results on image #12 from the STARE database, as shown in Fig. 2, demonstrates the progressive improvement achieved by the proposed method. While the B-COSFIRE-only approach (c) effectively captures major vessels but misses finer structures and exhibits fragmentation, and the phase-congruency-based method (d) shows higher sensitivity to thin vessels but introduces noticeable noise and false positives, the dual-feature combination without spatial constraint (e) strikes a better balance by integrating both structural and intensity information, yet still produces minor discontinuities. The proposed algorithm (f), incorporating dual features with spatial regularization, most closely approximates the expert annotations (b), successfully preserving continuity of major vessels, recovering more subtle vascular details, and effectively suppressing spurious noise, thereby yielding the most morphologically accurate and clinically reliable segmentation outcome.

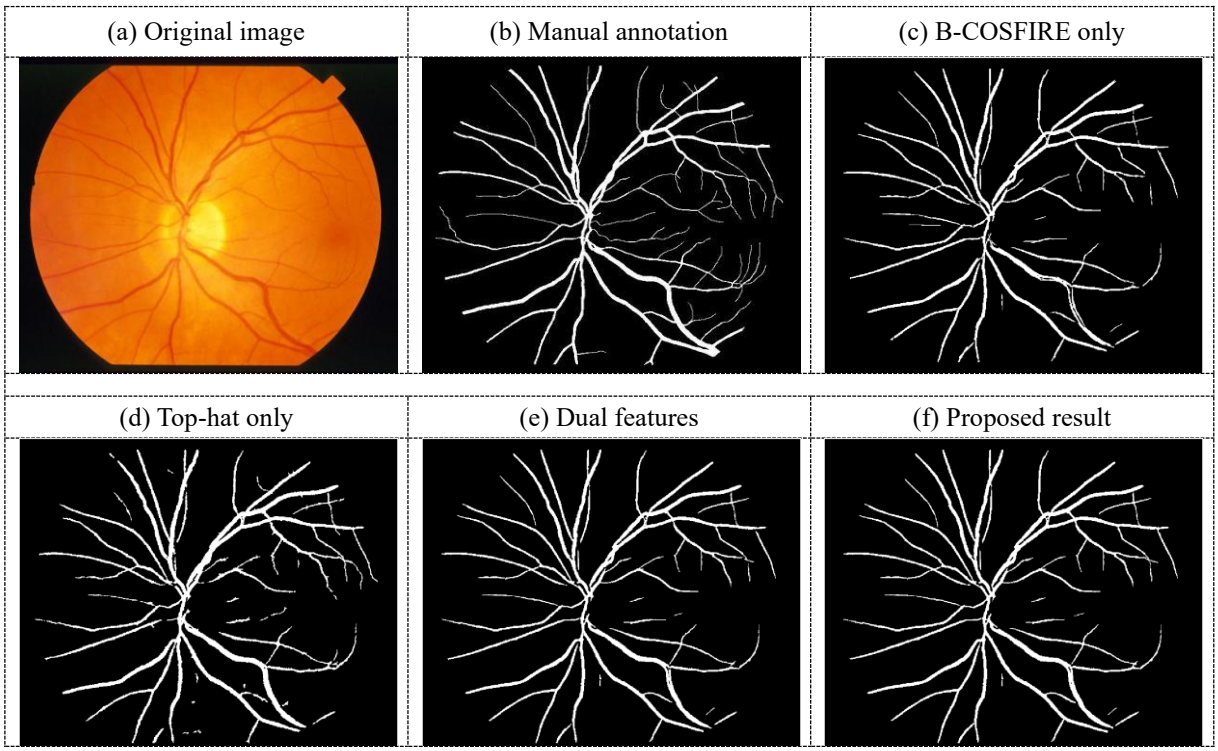

Figure 2. Comparative segmentation results on the STARE database. (a) Color fundus image; (b) Expert-annotated ground truth; (c) Retinal vessel segmentation result using only the B-COSFIRE feature; (d) Retinal

vessel segmentation result using only the Top-hat response; (e) Retinal vessel segmentation result based on dual features; (f) Segmentation result of the proposed method

3.3.3 Comparison with Other Methods

A quantitative comparative analysis was conducted between the proposed method and other unsupervised retinal vessel segmentation methods. The results are presented in Table 2.

Table 2. Performance comparison with other retinal vessel segmentation methods on the STARE database.

No.	Method	Acc	Se	Sp
1	Azzopardi et al.[9]	0.9497	0.7716	0.9701
2	Neto et al.[15]	0.8894	0.8344	0.9443
3	Fraz et al.[16]	0.9440	0.7311	0.9680
4	Yin et al.[17]	0.9325	0.8541	0.9419
5	Khomri et al.[18]	0.9400	0.7370	0.9620
6	Wang et al.[19]	0.9502	0.7654	0.9735
	Proposed method	0.9511	0.7439	0.9747

The proposed method achieves competitive performance compared to existing unsupervised approaches on the STARE database. It obtains the highest Accuracy and Specificity among all compared methods, demonstrating its overall robustness and exceptional ability to correctly identify background pixels, thereby minimizing false positives. While the Sensitivity is slightly lower than that of some methods (e.g., Neto et al.[15] and Yin et al.[17], which achieved 0.8344 and 0.8541, respectively), the proposed approach achieves a more favorable balance between Se and Sp. The superior Accuracy and Specificity indicate that the integration of dual features with spatial constraints effectively enhances segmentation precision and reliability.

4. Conclusion

The proposed algorithm effectively integrates the multi-directional and multi-scale vascular structural information obtained by B-COSFIRE filters with the local brightness enhancement provided by the Top-hat transformation, forming a complementary feature representation. By incorporating MRF-based spatial constraints into the information-theoretic co-clustering framework, the method successfully leverages pixel-neighborhood contextual information, thereby enhancing topological consistency. Experimental results on the STARE database demonstrate that the proposed retinal vessel segmentation approach, which combines dual-feature fusion and spatially constrained co-clustering, exhibits advantages in accuracy, robustness, and structural continuity. Both quantitative metrics and visual comparisons confirm significant improvements in the continuity of thin vessels and overall segmentation reliability. However, it was observed that the sensitivity for detecting fine vessels remains insufficient. Future work will focus on exploring novel co-clustering strategies to enhance the detection capability for subtle vascular structures.

Acknowledgments

The work in this paper was supported by the 2023 Annual Scientific Research Project of Jieyang Polytechnic (2023JYCKY06) and the 2023-2025 Innovation and Strengthening Higher Education Institutions Project of Jieyang Polytechnic (JYPCQKY23-0518).

References

[1] Liu, X., Tan, H., Wang, W., et al. (2024). Deep learning-based retinal vessel segmentation and hypertensive retinopathy quantification using a heterogeneous features cross-attention neural network. *Frontiers in Medicine, 11*, 1377479. https://doi.org/10.3389/fmed.2024.1377479

[2] Joseph, S., Selvaraj, J., Mani, I., et al. (2024). Diagnostic accuracy of artificial intelligence-based automated diabetic retinopathy screening in real-world settings: A systematic review and meta-analysis. *American Journal of Ophthalmology, 263*, 214-230. https://doi.org/10.1016/j.ajo.2024.02.012

[3] Qin, Q., & Chen, Y. (2024). A review of retinal vessel segmentation for fundus image analysis. *Engineering Applications of Artificial Intelligence, 128*, 107454. https://doi.org/10.1016/j.engappai.2023.107454

[4] Yin, N. B., Huang, M., & Liu, L. J. (2021). MS-UNet++: A retinal vessel segmentation method based on improved UNet++. *Journal of Optoelectronics Laser, 32*(1), 35-41.

https://doi.org/10.16136/j.joel.2021.01.0279

[5] Almarri, B., Kumar, B. N., Pai, H. A., et al. (2024). Redefining retinal vessel segmentation: Empowering advanced fundus image analysis with the potential of GANs. *Frontiers in Medicine, 11*, 1470941. https://doi.org/10.3389/fmed.2024.1470941

[6] Ouyang, J., Liu, S., Peng, H., et al. (2023). LEA U-Net: A U-Net-based deep learning framework with local feature enhancement and attention for retinal vessel segmentation. *SpringerPlus, 9*, 6753-6766. https://doi.org/10.1007/s40747-023-01095-3

[7] Chaudhuri, S., Chatterjee, S., Katz, N., et al. (1989). Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging, 8*(3), 263-269. https://doi.org/10.1109/42.34715

[8] Nandy, R. S., Chatterjee, R. K., & Das, A. (2024). Fundus vessel structure segmentation based on Bel-Hat transformation. *Microsystem Technologies, 30*(4), 439-453. https://doi.org/10.1007/s00542-023-05552-4

[9] Azzopardi, G., Strisciuglio, N., Vento, M., & Petkov, N. (2015). Trainable COSFIRE filters for vessel delineation with application to retinal images. *Medical Image Analysis, 19*(1), 46-57. https://doi.org/10.1016/j.media.2014.08.002

[10] Zheng, Q., Zheng, C. J., & Huang, M. X. (2025). Retinal vessel segmentation based on adaptive fuzzy local information C-means clustering. In *International Conference on Mechatronics and Intelligent Control (ICMIC 2024)* (Vol. 13447, pp. 447-455). SPIE. https://doi.org/10.1117/12.3052186

[11] Serra, J. (1982). *Image analysis and mathematical morphology*. Academic Press.

[12] Dhillon, I. S., Mallela, S., & Modha, D. S. (2003). Information-theoretic co-clustering. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 89-98). https://doi.org/10.1145/956750.956764

[13] Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 721-741. https://doi.org/10.1016/B978-0-08-051581-6.50057-X

[14] Hoover, A. D., Kouznetsova, V., & Goldbaum, M. (2000). Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging, 19*(3), 203-210. https://doi.org/10.1109/42.845178

[15] Neto, L. C., Ramalho, G. L., Neto, J. F. R., et al. (2017). An unsupervised coarse-to-fine algorithm for blood vessel segmentation in fundus images. *Expert Systems with Applications, 78*, 182-192. https://doi.org/10.1016/j.eswa.2017.02.015

[16] Fraz, M., Barman, S., Remagnino, P., et al. (2012). An approach to localize the retinal blood vessels using bit planes and centerline detection. *Computer Methods and Programs in Biomedicine, 108*(2), 600-616. https://doi.org/10.1016/j.cmpb.2011.08.009

[17] Yin, B., Li, H., Sheng, B., et al. (2015). Vessel extraction from non-fluorescein fundus images using orientation-aware detector. *Medical Image Analysis, 26*(1), 232-242. https://doi.org/10.1016/j.media.2015.09.002

[18] Khomri, B., Christodoulidis, A., Djerou, L., Babahenini, M. C., & Cheriet, F. (2018). Retinal vessel segmentation using the elite-guided multi-objective artificial bee colony algorithm. *IET Image Processing, 12*(12), 2163-2171. https://doi.org/10.1049/iet-ipr.2018.5425

[19] Wang, W., Wang, W., & Hu, Z. (2019). Retinal vessel segmentation approach based on corrected morphological transformation and fractal dimension. *IET Image Processing, 13*(13), 2538-2547. https://doi.org/10.1049/iet-ipr.2018.5636