UDC 004.93

MULTI-STAGE SEGMENTATION OF RETINAL LAYERS IN OCT B-SCANS

[1,2*]*Yusupov O.R.*, [2]*Shamsiyeva Kh.G.*, [2]*Badalova L.B.*

[+]ozodyusupov@gmail.com
[1]Samarkand State University named after Sharof Rashidov,
Samarkand, Uzbekistan.
[2]Digital technologies and artificial intelligence development research institute,
Tashkent, Uzbekistan

Abstract. This paper presents a training-free, interpretable, and computationally efficient multi-stage approach for segmenting retinal layers in OCT B-scans. The method is designed to remain robust under speckle noise, depth-dependent illumination nonuniformity, and device/protocol differences, and achieves practical speed on CPU. Clinically meaningful criteria – Dice, IoU, Boundary F1, and layer thickness error – are used for evaluation, alongside pragmatic guidance on parameter selection. The approach delivers reliable and repeatable results while maintaining a "quality–complexity–resource" balance suitable for resource-constrained environments.

Keywords: OCT; retinal layers; multi-stage segmentation; speckle noise; illumination nonuniformity; marker-controlled Watershed.

1 Introduction

Optical Coherence Tomography (OCT) is an advanced imaging technology that provides fast and non-invasive visualization of micron-scale cross sections of biological tissues [1]. It is widely used in ophthalmology for morphological analysis of retinal layers and in cardiology for assessing coronary artery wall conditions. Accurate segmentation of retinal layers enables automatic and repeatable computation of clinical indicators – thickness profiles, boundary smoothness, and the area and depth of pathological changes. These indicators are critical for early detection, monitoring, and treatment assessment of diseases such as glaucoma, diabetic retinopathy, and age-related macular degeneration (AMD).

Several physical properties of OCT images make segmentation challenging. Multiplicative speckle noise perturbs local intensities and attenuates edge responses. Attenuation of the signal with depth leads to illumination nonuniformity: deeper layers appear with reduced contrast and stronger shadowing. In many clinical scans, optical and motion artifacts occur, and there are differences between devices and scanning protocols. As a result, plain global thresholding or a single edge detector is insufficient to guarantee quality: layer boundaries can break, thin structures can fragment, or blend with background.

Modern deep-learning architectures (U-Net, SegNet, Attention U-Net) achieve high accuracy on OCT segmentation [2-4]. However, they require large, well-annotated datasets, substantial compute, and regular retuning. In clinical environments, annotation resources are often limited, scanner/protocol settings vary, and near-real-time processing is desirable. Especially in low-resource settings, training-free, interpretable, and parameter-stable solutions are preferred.

Carefully coordinated classical image processing methods can yield robust results on layered structures. CLAHE enhances local contrast without excessively amplifying speckle [5-6]. Percentile-calibrated hysteresis Canny produces stable edge maps across scans [7]. Morphological operations reconnect broken edges and remove spurious branches. The distance transform provides a strong signal for identifying "interior" points of layers. Marker-controlled Watershed balances edge and region information, separating segments in a physiologically plausible way [8-9]. With a prudent ordering and parameterization of these stages, practical quality can be achieved without large neural-model resources.

We provide a training-free, interpretable, CPU-efficient multi-stage pipeline for segmenting OCT B-scans into retinal layers, together with clear practical guidance. For each stage (contrast enhancement, edge extraction, morphological cleanup, marker design, Watershed), we explain parameter choices and their effect on segmentation quality. Clinically meaningful metrics – Dice, IoU, Boundary F1, and layer thickness error (e.g., Mean Absolute Thickness Error, MATE) – are used for evaluation. The approach is robust on CPU without supervised data, tolerant to scanner and protocol shifts, and provides actionable tuning tips for failure cases.

While classical multi-stage methods cannot always replace deep learning, especially in heavily artifacted cases or multimorbidity, they provide reliable and repeatable results under resource constraints while preserving the "quality–complexity–resource" balance. Future improvements include speckle-aware diffusion (SRAD) or non-local filtering and hybridization with lightweight learnable modules [6].

Organization. We first state the problem and image model; then describe each stage of the multi-stage pipeline with parameter selection guidance; next, we define the evaluation protocol and ablation study; finally, we discuss practical relevance, limitations, and future directions.

2 Problem statement (OCT layer segmentation)

Let the OCT B-scan image domain be $\Omega \subset \mathbb{R}^2$ and the input grayscale image $I: \Omega \to [0, L]$. A common physical observation model is

$$I(x,y) \approx L(x,y)\, I^*(x,y)\, \eta(x,y), \tag{1}$$

where $L(x,y)$ is slowly varying illumination/attenuation, $\eta(x,y) > 0$ is multiplicative (speckle) noise, and I^* denotes the "ideal" noise-free signal. The goal is to partition the image into n anatomical layers $S = \{S_1, \dots, S_n\}$ and recover their boundaries

$$\Gamma_i = \{(x, y_i(x))\}, \quad i = 1, \dots, n-1, \tag{2}$$

such that $S_i \cap S_j = \varnothing$ for $i \neq j$, $\bigcup_{i=1}^{n} S_i = \Omega$, and the retinal column-wise order holds:

$$\forall x: \quad y_1(x) < y_2(x) < \cdots < y_{n-1}(x). \tag{3}$$

As a practical solution, we form a local cost (relief) using the gradient magnitude of a contrast-enhanced image I_c,

$$G(x,y) = \| \nabla I_c(x,y) \|, \tag{4}$$

and a distance transform $D(x,y)$ computed on a morphologically cleaned edge mask. The composite potential is

$$\varphi(x,y) = D(x,y) + \lambda\, G(x,y), \qquad \lambda > 0. \tag{5}$$

Given a set of foreground markers F, the geodesic distance for each $p \in \Omega$ is

$$\mathrm{dist}_{\varphi}(p,q) = \inf_{\gamma(0)=p, \gamma(1)=q} \int_0^1 \varphi(\gamma(t)) \, \| \gamma'(t) \| \, dt, \qquad q \in F, \tag{6}$$

and we assign the "cheapest" marker by

$$W(p) = \operatorname*{argmin}_{q \in F} \mathrm{dist}_{\varphi}(p,q). \tag{7}$$

This produces regions that are subsequently reconciled with column-wise ordering and anatomical priors, optionally with projection/smoothing constraints. The output is either the label map $\hat{L} \in \{1, \dots, n\}^{H \times W}$ (a layer label per pixel) or the set of boundaries $\{\Gamma_i\}$. Reliable layer thickness profiles are then obtained as

$$T_i(x) = y_i(x) - y_{i-1}(x), \qquad i = 1, \dots, n-1. \tag{8}$$

3 An approach to solving the problem

The proposed solution follows a classical multi-stage pipeline designed for retinal layer segmentation in OCT B-scan images. The process sequentially performs contrast stabilization, edge extraction, morphological refinement, distance transformation with marker generation, marker-controlled watershed segmentation, and final post-processing. Each stage progressively refines boundary localization while maintaining computational efficiency and robustness.

The preprocessing stage performs affine normalization of intensity values to stabilize contrast across scans and improve the transferability of parameters. Median filtering, being a nonlinear order-statistic filter, effectively suppresses impulsive noise and speckle-related outliers while preserving edges better than Gaussian smoothing. This ensures that subsequent enhancement and edge detection stages operate on a stable, denoised input. The normalization and denoising step can be expressed as

$$I_0(x,y) = \frac{I(x,y) - \min\limits_{\Omega} I}{\max\limits_{\Omega} I - \min\limits_{\Omega} I} \cdot L, \qquad I_1 = \mathrm{Median}(I_0; 3 \times 3). \tag{9}$$

Local contrast enhancement is achieved using Contrast-Limited Adaptive Histogram Equalization (CLAHE), which compensates for depth-dependent attenuation typical of OCT data. CLAHE equalizes local histograms with clipped saturation, enhancing micro-contrast without overamplifying global brightness variations. Tiled processing with bilinear interpolation preserves smooth transitions between adjacent tiles and prevents block artifacts:

$$I_c = \mathrm{CLAHE}(I_1; \mathrm{clipLimit}, \mathrm{tileGridSize}), \tag{10}$$

where $\mathrm{tileGridSize} \in \{8 \times 8, 12 \times 12\}$ and $\mathrm{clipLimit} \in [2.0,3.0]$.

Edge detection is performed using an adaptive Canny operator. Gradients are computed in horizontal and vertical directions, followed by non-maximum suppression and double-threshold hysteresis. The high threshold is set adaptively using an image percentile $p \in [0.7,0.9]$, while the low threshold is defined as a fraction of the high threshold ($\alpha \in [0.3,0.5]$):

$$I_x = I_c * S_x, \qquad I_y = I_c * S_y, \qquad G(x,y) = \sqrt{I_x^2(x,y) + I_y^2(x,y)}, \tag{11}$$

$$T_h = \mathrm{quantile}(G; p), \qquad T_l = \alpha T_h. \tag{12}$$

Strong and weak edge sets are defined as

$$S = \{(x,y): G(x,y) \geq T_h\}, \qquad W = \{(x,y): T_l \leq G(x,y) < T_h\}, \tag{13}$$

and the final edge map E is obtained through hysteresis thresholding [10]:

$$E = S \cup \{w \in W \mid \exists \mathrm{apath} w \rightsquigarrow s, s \in S\}. \tag{14}$$

Morphological cleanup follows to geometrically regularize the detected edges. Opening removes small spurious components, and closing bridges minor gaps or discontinuities. Since retinal layers are predominantly horizontal, structuring elements are chosen as small disks or short horizontal lines:

$$M = (E \circ B) \bullet B, \tag{15}$$

where B is a structuring element (disk of radius $r \in \{1,2\}$ or a horizontal line of 3–5 pixels), and $\circ$ and $\bullet$ denote opening and closing, respectively.

On the complement of the edge mask, the Euclidean distance transform encodes the distance of each pixel from the nearest edge. The smoothed distance map is used to identify regional maxima as foreground markers, while background markers stabilize the outer regions:

$$D(x,y) = \min_{(x\prime,y\prime): M(x\prime,y\prime)=1} \sqrt{(x-x')^2 + (y-y')^2}, \tag{16}$$

$$\widetilde{D} = \mathrm{Gaussian}(D; \sigma = 1), \qquad F = \mathrm{RegMax}(\widetilde{D}), \qquad B = \mathrm{BackgroundMarkers}(M).$$

To integrate region and boundary information, a local relief function is defined as

$$\varphi(x,y) = D(x,y) + \lambda \parallel \nabla I_c(x,y) \parallel, \qquad \lambda > 0, \tag{17}$$

which balances adherence to strong edges and regional smoothness. The geodesic distance under φ between points p and q is defined as

$$\mathrm{dist}_\varphi(p,q) = \inf_{\gamma(0)=p,\gamma(1)=q} \int_0^1 \varphi(\gamma(t)) \parallel \gamma'(t) \parallel \, dt. \tag{18}$$

Larger values of λ encourage sharper boundary alignment, while smaller λ promotes smoother segmentation.

Marker-controlled watershed segmentation is applied to the relief surface. In this mode, initial minima are defined by the foreground and background markers, which greatly reduce over-segmentation. The pixel-wise assignment is given by

$$W(p) = \operatorname*{argmin}_{q \in F} \mathrm{dist}_\varphi(p,q), \qquad p \in \Omega, \tag{19}$$

producing a label map

$$\hat{L} \in \{1, \dots, n\}^{H \times W}.$$

Post-processing involves removing small connected components and smoothing extracted boundaries column-wise using median or spline filtering. The retinal layer order constraint is then enforced to ensure anatomical validity:

$$\hat{L} \leftarrow \mathrm{RemoveSmallCC}(\hat{L}; \mathrm{minArea}), \qquad \Gamma_i = \{(x, y_i(x))\}_{x=1}^{W}, \tag{20}$$

$$\forall x: \; y_1(x) < y_2(x) < \cdots < y_{n-1}(x).$$

The entire process can be interpreted from a variational perspective as the minimization of an energy functional combining boundary fidelity, smoothness, and layer order regularization:

$$E(L) = D_{\mathrm{bnd}}(L; I_c) + \alpha R_{\mathrm{smooth}}(L) + \beta R_{\mathrm{order}}(L) \to \min, \tag{21}$$

where

$$D_{\mathrm{bnd}}(L; I_c) \approx -\sum_{(x,y)\in\Omega} \mathbf{1}_{\partial L}(x,y) \parallel \nabla I_c(x,y) \parallel, \qquad R_{\mathrm{smooth}}(L) = \sum_{(x,y)} \parallel \nabla L(x,y) \parallel.$$

The marker-controlled watershed provides a discrete and computationally efficient approximation to minimizing this energy.

The final output consists of a labeled segmentation map $\hat{L} \in \{1, \dots, n\}^{H \times W}$ or boundary curves $\{\Gamma_i\}_{i=1}^{n-1}$, from which layer thickness profiles are derived:

$$T_i(x) = y_i(x) - y_{i-1}(x), \qquad i = 1, \dots, n-1. \tag{22}$$

These thickness profiles are subsequently used for clinical evaluation, such as mean or median layer thickness estimation and identification of anomalous zones associated with pathologies.

Each component of the pipeline can be implemented with near-linear computational complexity $O(N)$, where $N = H \cdot W$. Algorithmic stability depends on maintaining a lower-bounded relief $\varphi \geq c > 0$, a sufficient number of markers, and effective morphological cleanup. Proper selection of parameter ranges, including CLAHE tile size, Canny percentile thresholds, and λ, supports cross-scanner generalization and ensures reliable segmentation in diverse clinical imaging conditions.

4 Experimental evaluation

The practical effectiveness of the proposed classical multi-stage segmentation pipeline was evaluated under a training-free setup, where parameters were chosen once on a small validation subset and then fixed for all test scans. Three data domains were considered: (i) Set A containing standard retinal B-scans with moderate signal-to-noise ratio and minimal artifacts, (ii) Set B consisting of artifact-heavy B-scans affected by shadowing, saturation, and motion distortions, and (iii) Set C representing a domain-shift scenario with scans obtained from a different imaging device or

acquisition protocol. For each dataset, retinal layer boundaries $y_i^{gt}(x)$ for $i = 1, \dots, n-1$ were annotated by expert clinicians as one-dimensional polylines, and in cases of multiple annotators, consensus or mean annotations were used. The data were split into validation (10–20%) and test (80–90%) subsets, while no training split was required since the approach is fully training-free.

Evaluation metrics included region-based and boundary-based measures. Dice and IoU scores were computed as [11,12]

$$\text{Dice} = \frac{2|S \cap G|}{|S|+|G|}, \qquad \text{IoU} = \frac{|S \cap G|}{|S \cup G|},$$

respectively. Boundary accuracy was quantified using the Boundary-F1 (BF1) metric within a tolerance of typically two pixels,

$$\text{BF1} = \frac{2 \cdot \text{Prec} \cdot \text{Rec}}{\text{Prec}+\text{Rec}}.$$

To assess layer thickness fidelity, predicted and reference thickness profiles

$$T_i^{pred}(x) = y_i^{pred}(x) - y_{i-1}^{pred}(x), \qquad T_i^{gt}(x) = y_i^{gt}(x) - y_{i-1}^{gt}(x)$$

were compared using the mean absolute error

$$\text{MAE}_T = \frac{1}{\sum_i |X_i|} \sum_{i=1}^{n-1} \sum_{x \in X_i} |T_i^{pred}(x) - T_i^{gt}(x)|.$$

Additionally, one-dimensional structural similarity (SSIM) computed column-wise and the smoothness of layer curves (RMS of second differences) were reported to capture geometric regularity.

Comparisons were made against several classical baselines: (B1) global Otsu thresholding [13] followed by morphology, (B2) Canny edge detection [14] with plain Watershed segmentation without markers, (B3) a gradient-weighted GraphCut formulation [15], and (B4) variational partitioning with total variation regularization. All baselines and the proposed method were tuned using the same validation-based parameter selection strategy. The proposed pipeline follows the sequence: Normalize → Median → CLAHE → Canny → Morphology → Distance Transform and Markers → Marker-Controlled Watershed → Post-Processing.

All experiments were executed on an x86_64 CPU platform. Image intensities were normalized to $[0, L]$ with $L = 255$. CLAHE parameters included tile grid sizes of 8×8 or 12×12 and clip limits between 2.0 and 3.0. For adaptive Canny, parameters were set to $p = 0.8$ and $\alpha = 0.4$. Morphological operations employed disk structuring elements of radius 1–2 or horizontal lines of 3–5 pixels. Distance transform smoothing used $\sigma = 1$, with the relief-balance weight $\lambda \in [0.5,2.0]$ (default 1.0), and a minimum area threshold between 0.05% and 0.1% of total pixels. Parameters were optimized through grid or random search over clipLimit $\in \{2.0,2.5,3.0\}$, $p \in \{0.75,0.80,0.85\}$, $\alpha \in \{0.3,0.4,0.5\}$, and $\lambda \in \{0.5,1.0,1.5,2.0\}$. The Pareto-optimal configuration was determined based on BF1 and MAE_T trade-offs and transferred unchanged to the test phase.

Ablation experiments were conducted to assess the contribution of individual stages. Removing CLAHE led to weaker gradients and drops in BF1 and IoU; omitting markers caused pronounced over-segmentation and larger MAE_T; excluding morphological refinement resulted in discontinuous and rougher boundaries; and skipping distance-transform smoothing introduced unstable seeds and excessive small fragments. Each ablated variant was evaluated over all datasets with mean $\pm$ std values for Dice, IoU, BF1, and MAE_T.

Statistical analysis employed the Wilcoxon signed-rank test with $\alpha = 0.05$ to assess significance of paired differences, testing the null hypothesis that median differences equal zero. p-values were reported alongside 95% bootstrap confidence intervals (1,000–5,000 resamples) for mean and median metrics [16,17]. Computational profiling indicated per-stage runtime scaling approximately as

$$t_{\text{CLAHE}} + t_{\text{Canny}} + t_{\text{Morph}} + t_{\text{DT/markers}} + t_{\text{WS}} + t_{\text{Post}} \approx O(HW),$$

with total processing time on the order of seconds for 1-megapixel images and a memory footprint of only a few tens of megabytes. CLAHE and morphological operations can be parallelized across CPU threads, and no GPU acceleration is required.

Qualitative visualizations include the raw input image, contrast-enhanced version, Canny edges, morphological mask, smoothed distance map with markers, and final segmentation with boundary overlays, with artifact regions highlighted for clarity. Quantitative results are summarized as mean $\pm$ std tables supplemented with violin and box plots. Sensitivity plots showing BF1 versus λ and MAE_T versus clipLimit illustrate trade-offs and stability regions.

Robustness evaluation shows that parameters generalize well to unseen domains: on the domain-shifted Set C, performance remains stable without re-tuning, while on artifact-heavy Set B, CLAHE and morphology notably improve segmentation coherence. In challenging cases involving severe shadows or saturation, reducing λ and reinforcing morphological closing help stabilize boundaries; for very thin layers, increasing distance-transform smoothing or lowering the minArea threshold proves effective.

To ensure reproducibility, the release package includes the full configuration (clipLimit, tileGridSize, p, α, structuring element type, λ, minArea, σ), deterministic random seeds, library version details, intermediate outputs of all pipeline stages, and scripts for reporting Dice, IoU, BF1, MAE_T, and statistical significance tests.

5 Results

We compare Ours against baselines B1–B4. All parameters are selected once on validation and transferred unchanged to test. For axial scaling we assume 1 pixel $\approx$ 3.9 μm.

5.1 Quantitative results (main metrics)

Table 1 reports test metrics (mean$\pm$std). Ours outperforms baselines in boundary accuracy (BF1), region overlap (Dice/IoU), and layer thickness error (MAE_T).

Tab. 1. Illustrative results (values representative of ranges in the literature)

Method	Dice	IoU	BF1	MAE_T (px / μm)	Time (s/img)
B1: Otsu + Morph	0.79$\pm$0.07	0.66$\pm$0.08	0.75$\pm$0.09	5.4$\pm$2.0 / 21.1$\pm$7.8	0.62
B2: Canny + Plain WS	0.82$\pm$0.06	0.70$\pm$0.07	0.80$\pm$0.08	4.8$\pm$1.8 / 18.7$\pm$7.0	0.68
B3: GraphCut-variant	0.84$\pm$0.05	0.73$\pm$0.06	0.82$\pm$0.07	4.3$\pm$1.6 / 16.8$\pm$6.2	0.95
B4: TV-Partition	0.85$\pm$0.05	0.74$\pm$0.06	0.83$\pm$0.06	4.0$\pm$1.5 / 15.6$\pm$5.9	1.10
Ours (full)	0.89$\pm$0.04	0.80$\pm$0.05	0.88$\pm$0.05	3.2$\pm$1.2 / 12.5$\pm$4.7	0.80

Per-layer thickness errors are summarized in Table 2.

Tab. 2. Per-layer thickness error statistics (1 px ≈ 3.9 μm)

Layer	**MAE** (px)	**MAE** (μm)	**Median** (μm)	**IQR** (μm)
T_1 (ILM–RNFL)	3.0	11.7	10.9	5.3
T_2 (RNFL–GCL)	3.4	13.3	12.5	5.9
T_3 (GCL–IPL)	2.8	10.9	10.2	5.0
T_4 (IPL–INL)	3.6	14.0	13.1	6.2
T_5 (INL–OPL)	3.1	12.1	11.4	5.6
T_6 (OPL–ONL)	3.5	13.7	12.9	6.0

5.2 Ablation study

Table 3 shows metric changes when removing components. The biggest drops occur without markers and without CLAHE.

Tab. 3. Ablation: effect of removing stages

Variant	**Dice**	**IoU**	**BF1**	MAE_T (px)
Full chain (Ours)	**0.89**	**0.80**	**0.88**	**3.2**
– CLAHE	0.86	0.76	0.84	3.9
– Morph	0.87	0.78	0.85	3.6
– Markers (plain WS)	0.83	0.71	0.81	4.6
– DT smoothing	0.88	0.79	0.86	3.5

5.3 Statistical significance

Table 4 reports paired Wilcoxon signed-rank test ($\alpha = 0.05$) p-values comparing Ours to each baseline; most improvements are significant.

Tab. 4. Wilcoxon signed-rank test ($\alpha = 0.05$)

Comparison	p_{Dice}	p_{IoU}	p_{BF1}	p_{MAE_T}
Ours vs B1	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$	$< 10^{-4}$
Ours vs B2	3.0×10^{-4}	4.0×10^{-4}	5.0×10^{-4}	7.0×10^{-4}
Ours vs B3	1.1×10^{-3}	1.6×10^{-3}	9.0×10^{-4}	1.8×10^{-3}
Ours vs B4	1.8×10^{-2}	2.1×10^{-2}	9.0×10^{-3}	1.2×10^{-2}

5.4 Time and resources

For a 1 MP image, per-stage average times are summarized in Table 5; total ≈ 0.8 s/img (CPU), no GPU required.

Tab. 5. Stage-wise runtime profile (CPU, 1 MP)

	CLAHE	Grad/Canny	Morph	DT/Markers	WS	Post	Total
Time (s/img)	0.18	0.12	0.10	0.15	0.18	0.07	**0.80**

6 Discussion

Ours improves boundary accuracy (BF1 ↑) and reduces thickness error (MAE_T ↓) vs baselines. Gains arise from: (i) CLAHE ⇒ stabilized gradients; (ii) adaptive Canny ⇒ cross-scan calibration; (iii) DT seeds ⇒ less over-segmentation; (iv) balanced relief D–∇I_c.

7 Conclusion

In this study, a training-free classical multi-stage pipeline for OCT B-scan segmentation was proposed and theoretically justified. The pipeline consists of the following sequence of operations: local contrast enhancement using CLAHE, adaptive Canny edge detection, morphological refinement, distance transform with markers, marker-controlled watershed segmentation, and final post-processing. The proposed method fuses regional and edge-based information through the integrated cost function

$$\varphi(x, y) = D(x, y) + \lambda \parallel \nabla I_c(x, y) \parallel$$

which ensures coherent and stable boundary localization.

Experimental results demonstrate that the proposed approach consistently outperforms classical baselines in terms of $BF1$, $Dice/IoU$, and MAE_T metrics. The improvements in accuracy are primarily due to gradient stabilization achieved by CLAHE, reduced over-segmentation through marker control, and balanced relief representation. Once parameters are selected on a validation set, they generalize effectively across different domains, including datasets affected by artifacts, varying imaging devices, and acquisition protocols. The proposed method is also computationally efficient, with approximately $O(HW)$ complexity per stage, requiring about 0.8 seconds to process a 1-megapixel image on a CPU, and does not rely on GPU acceleration.

Nevertheless, some limitations remain. Strong shadows or saturated regions can generate false edges, and very thin layers or extremely low SNR may lead to sparse or missing markers. In such cases, globally fixed parameters may need local adaptation to maintain performance. Future work aims to incorporate speckle-aware denoising methods such as SRAD or NLM, introduce spatially adaptive weighting for λ, and apply 1D spline or Kalman-based smoothing with uncertainty estimation. Further extensions will explore 3D C-scan temporal and cross-slice regularization, multimodal OCT/OCTA fusion, and semi-automatic or active-learning-based marker optimization for adaptive tuning.

In summary, the proposed classical pipeline provides accurate, robust, and computationally efficient OCT segmentation without the need for large annotated datasets. It is particularly well-suited for low-resource clinical environments and offers a solid foundation for future developments involving 3D integration, adaptive parameterization, and uncertainty-aware segmentation frameworks.

References

[1] *Huang D., Swanson E.A., Lin C.P., Schuman J.S., Stinson W.G., Chang W., Hee M.R., Flotte T., Gregory K., Puliafito C.A., Fujimoto J.G.* Optical coherence tomography // Science. – 1991. – Vol. 254, No. 5035. – P. 1178–1181.

[2] *Ronneberger O., Fischer P., Brox T.* U-net: Convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention (MICCAI). – 2015. – Vol. 9351. – P. 234–241.

[3] *Badrinarayanan V., Kendall A., Cipolla R.* Segnet: A deep convolutional encoder-decoder architecture for image segmentation // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2017. – Vol. 39, No. 12. – P. 2481–2495.

[4] *Oktay O., Schlemper J., Le Folgoc L., Lee M., Heinrich M., Misawa K., Mori K., McDonagh S., Hammerla N.Y., Kainz B., Glocker B., Rueckert D.* Attention u-net: Learning where to look for the pancreas // arXiv preprint arXiv:1804.03999. – 2018.

[5] *Zuiderveld K.* Contrast limited adaptive histogram equalization // Graphics Gems IV / ed. Heckbert P.S. – San Diego: Academic Press, 1994. – P. 474–485.

[6] *Yu Y., Acton S.T.* Speckle reducing anisotropic diffusion // IEEE Transactions on Image Processing. – 2002. – Vol. 11, No. 11. – P. 1260–1270.

[7] *Canny J.* A computational approach to edge detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1986. – Vol. 8, No. 6. – P. 679–698.

[8] *Vincent L., Soille P.* Watersheds in digital spaces: An efficient algorithm based on immersion simulations // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1991. – Vol. 13, No. 6. – P. 583–598.

[9] *Arbelaez P., Maire M., Fowlkes C., Malik J.* Contour detection and hierarchical image segmentation // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2011. – Vol. 33, No. 5. – P. 898–916.

[10] *Otsu N.* A threshold selection method from gray-level histograms // IEEE Transactions on Systems, Man, and Cybernetics. – 1979. – Vol. 9, No. 1. – P. 62–66.

[11] *Dice L.R.* Measures of the amount of ecologic association between species // Ecology. – 1945. – Vol. 26, No. 3. – P. 297–302.

[12] *Sørensen T.* A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons // Biologiske Skrifter / Kongelige Danske Videnskabernes Selskab. – 1948. – Vol. 5. – P. 1–34.

[13] *Boykov Y.Y., Jolly M.-P.* Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images // Proceedings of IEEE International Conference on Computer Vision (ICCV). – Vancouver, Canada, 2001. – Vol. 1. – P. 105–112.

[14] *Chan T.F., Vese L.A.* Active contours without edges // IEEE Transactions on Image Processing. – 2001. – Vol. 10, No. 2. – P. 266–277.

[15] *Rudin L.I., Osher S., Fatemi E.* Nonlinear total variation based noise removal algorithms // Physica D: Nonlinear Phenomena. – 1992. – Vol. 60, No. 1-4. – P. 259–268.

[16] *Wilcoxon F.* Individual comparisons by ranking methods // Biometrics Bulletin. – 1945. – Vol. 1, No. 6. – P. 80–83.

[17] *Efron B.* Bootstrap methods: Another look at the jackknife // The Annals of Statistics. – 1979. – Vol. 7, No. 1. – P. 1–26.

МНОГОЭТАПНАЯ СЕГМЕНТАЦИЯ СЛОЁВ СЕТЧАТКИ НА ОКТ B-SCAN

[1,2*]***Юсупов О.Р.,*** [2]***Шамсиева Х.Г.,*** [2]***Бадалова Л.Б.***

[+]ozodyusupov@gmail.com
[1]Самаркандский государственный университет имени Шарофа Рашидова, Самарканд, Узбекистан
[2] Научно-исследовательский институт развития цифровых технологий и искусственного интеллекта, Ташкент, Узбекистан

Аннотация. В данной работе представлен многоэтапный подход к сегментации слоёв сетчатки на ОКТ B-scan, не требующий обучения, интерпретируемый и вычислительно эффективный. Метод разработан таким образом, чтобы сохранять устойчивость к спекл-шуму, глубинно-зависимой неравномерности освещения и межустройственным/межпротокольным различиям, обеспечивая при этом практичную скорость выполнения на CPU. Для оценки применяются клинически значимые метрики — Dice, IoU, Boundary F1, а также ошибка определения толщины слоёв — в сочетании с практическими рекомендациями по выбору параметров. Предложенный подход обеспечивает надёжные и воспроизводимые результаты, сохраняя баланс «качество–сложность–ресурсы», что делает его применимым в условиях ограниченных вычислительных возможностей.

Ключевые слова: ОКТ; слои сетчатки; многоэтапная сегментация; спекл-шум; неравномерность освещения; маркер-управляемый Watershed.

OKT B-SCAN TASVIRLARIDA TO‘R PARDA QATLAMLARIDA KO‘P BOSQICHLI SEGMENTATSIYA

[1,2*]***Yusupov O.R.,*** [2]***Shamsiyeva X.G.,*** [2]***Badalova L.B.***

[+]ozodyusupov@gmail.com
[1]Sharof Rashidov nomidagi Samarqand davlat universiteti, Samarqand, O‘zbekistan.
[2]Raqamli texnologiyalar va sun’iy intellektni rivojlantirish ilmiy-tadqiqot instituti, Toshkent, O‘zbekistan.

Annotatsiya. Ushbu maqolada OKT B-scan tasvirlarida to‘r parda qatlamlarini segmentatsiya qilish uchun o‘qitish talab etilmaydigan, interpretatsiya qilish oson va hisoblash resurslari bo‘yicha samarali bo‘lgan ko‘p bosqichli yondashuv taqdim etiladi. Taklif etilgan usul dog‘li shovqini, chuqurlikka bog‘liq yoritilishning notekisligi hamda qurilmalar/protokollar o‘rtasidagi farqlar sharoitida barqaror ishlashi uchun ishlab chiqilgan bo‘lib, CPU muhitida ham amaliy tezlikni ta’minlaydi. Baholashda klinik ahamiyatga ega ko‘rsatkichlar — Dice, IoU, Boundary F1 hamda qatlam qalinligini aniqlash xatosi — qo‘llaniladi, shuningdek parametrlarni tanlash bo‘yicha amaliy tavsiyalar beriladi. Mazkur yondashuv resurslar cheklangan sharoitlarda ham “sifat–murakkablik–resurs” muvozanatini saqlagan holda ishonchli va takrorlanadigan natijalarni ta’minlaydi.

Kalit so‘zlar: OKT; to‘r parda qatlamlari; ko‘p bosqichli segmentatsiya; dog‘li shovqin; yoritilishning notekisligi; marker-boshqaruvli Watershed.