A MULTI-STAGE APPROACH FOR RETINAL LAYER SEGMENTATION IN OCT B-SCAN IMAGES

Shamsieva Kh.
PhD student, Digital technologies and artificial intelligence development research institute, 17A, Boz-2,
Tashkent, 100125,
Uzbekistan
shamsiyeva117@gmail.com

Annotation: This study introduces an interpretable, training-free, and computationally efficient multi-stage framework for the segmentation of retinal layers in OCT B-scan images. The proposed approach is robust to common challenges such as speckle noise, depth-related illumination inconsistencies, and variations across imaging devices and acquisition protocols, while maintaining fast execution on standard

CPU systems. Evaluation is carried out using clinically relevant metrics, including Dice coefficient, IoU, Boundary F1-score, and layer thickness error, along with practical recommendations for parameter tuning. The method demonstrates consistent and reliable performance, achieving an effective balance between quality, computational complexity, and resource consumption, making it suitable for environments with limited computational capacity.

Keywords: OCT, retinal layer segmentation, multi-stage approach, speckle noise, illumination variation, marker-controlled Watershed.

МНОГОЭТАПНЫЙ ПОДХОД К СЕГМЕНТАЦИИ СЛОЁВ СЕТЧАТКИ НА ОКТ В-СКАНАХ

Шамсиева Х.Г.
докторант, НИИ Развития цифровых технологий и искусственного интеллекта, 100125, г.Ташкент, Буз-2, Республика Узбекистан
shamsiyeva117@gmail.com

Аннотация: Аннотация: В данной работе представлен интерпретируемый, не требующий обучения и вычислительно эффективный многоэтапный подход к сегментации слоёв сетчатки на ОКТ В-сканах. Предложенный метод устойчив к распространённым проблемам, таким как спекл-шум, неоднородность освещения, зависящая от глубины, а также различия между устройствами визуализации и протоколами получения данных, при этом обеспечивая высокую скорость выполнения на стандартных CPU-системах. Оценка проводится с использованием клинически значимых метрик, включая коэффициент Дайса, IoU, Boundary F1-меру и ошибку толщины слоя, а также даются практические рекомендации по настройке параметров. Метод демонстрирует стабильные и надёжные результаты, достигая эффективного баланса между качеством, вычислительной сложностью и потреблением ресурсов, что делает его подходящим для условий с ограниченными вычислительными возможностями.

Ключевые слова: ОКТ, сегментация слоёв сетчатки, многоэтапный подход, спекл-шум, вариации освещения, маркер-управляемый Watershed.

OKT B-SKAN TASVIRLARIDA TO‘R PARDA QATLAMLARINI KO‘P BOSQICHLI SEGMENTATSIYA QILISH YONDASHUVI

Shamsiyeva X.G.
doktorant, PhD, Raqamli texnologiyalar va sun’iy intellektni rivojlantirish ilmiy-tadqiqot institute shamsiyeva117@gmail.com

Annotatsiya: Ushbu tadqiqotda OKT B-skan tasvirlarida to‘r parda qatlamlarini segmentatsiya qilish uchun interpretatsiya qilinadigan, o‘qitishni talab qilmaydigan va hisoblash jihatdan samarali ko‘p bosqichli yondashuv taklif etiladi. Taklif etilgan usul spekl shovqin, chuqurlikka bog‘liq yoritilish notekisligi hamda tasvir olish qurilmalari va protokollari o‘rtasidagi farqlar kabi keng tarqalgan muammolarga nisbatan barqaror bo‘lib, shu bilan birga oddiy CPU tizimlarida tez ishlashni ta’minlaydi. Baholash klinik jihatdan muhim bo‘lgan ko‘rsatkichlar — Dice koeffitsienti, IoU, Boundary F1 o‘lchovi va qatlam qalinligi xatosi asosida amalga oshiriladi, shuningdek parametrlarni sozlash bo‘yicha amaliy tavsiyalar beriladi. Taklif etilgan usul barqaror va ishonchli natijalarni ko‘rsatib, sifat, hisoblash murakkabligi va resurs sarfi o‘rtasida samarali muvozanatni ta’minlaydi, bu esa uni hisoblash imkoniyatlari cheklangan muhitlarda qo‘llashga mos qiladi.

Kalit so‘zlar: OKT, to‘r parda qatlamlari segmentatsiyasi, ko‘p bosqichli yondashuv, spekl shovqin, yoritilish variatsiyasi, marker-boshqaruvli Watershed.

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.

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

$$I(x, y) \approx L(x, y)\, I^*(x, y)\, \eta(x, y), \quad (1)$$

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

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

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:

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

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

$$G(x, y) = \| \nabla I_c(x, y) \|, \quad (4)$$

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

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

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

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

and we assign the "cheapest" marker by

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

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 $L \in \{1, \dots, n\}^{H \times W}$ (a layer label per pixel) or the set of boundaries $\{\Gamma_i\}$. Reliable layer thickness profiles are then obtained as

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

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.

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.

$$I_0(x, y) = \frac{I(x, y) - \min_\Omega I}{\max_\Omega I - \min_\Omega I} \cdot L, \qquad I = \mathrm{Median}(I_0; 3 \times 3). \quad (9)$$

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:

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

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. 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 ($\alpha \in [0.3,0.5]$):

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

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

Strong and weak edge sets are defined as

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

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

$$E = S \cup \{w \in W \mid \exists \mathrm{path} w \rightsquigarrow s, s \in S\}. \quad (14)$$

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:

$$M = (E \circ B) \bullet B, \quad (15)$$

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:

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

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

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

$$\varphi(x,y) = D(x,y) + \lambda \parallel \nabla I_c(x,y) \parallel, \quad \lambda > 0, \quad (17)$$

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

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

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

$$W(p) = \underset{q\in F}{argmin} dist_\varphi(p,q), \quad p \in \Omega, \quad (19)$$

producing a label map $L \in \{1,\dots,n\}^{H\times W}$.

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:

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

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

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

Experimental evaluation. 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 $\mathrm{clipLimit} \in \{2.0, 2.5, 3.0\}$, $p \in \{0.75, 0.80, 0.85\}$, $\alpha \in \{0.3, 0.4, 0.5\}$, 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.

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 ≈ 3.9 μm.

Table 1 reports test metrics (mean±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±0.07	0.66±0.08	0.75±0.09	5.4±2.0 / 21.1±7.8	0.62
B2: Canny + Plain WS	0.82±0.06	0.70±0.07	0.80±0.08	4.8±1.8 / 18.7±7.0	0.68
B3: GraphCut-variant	0.84±0.05	0.73±0.06	0.82±0.07	4.3±1.6 / 16.8±6.2	0.95
B4: TV-Partition	0.85±0.05	0.74±0.06	0.83±0.06	4.0±1.5 / 15.6±5.9	1.10
Ours (full)	0.89±0.04	0.80±0.05	0.88±0.05	3.2±1.2 / 12.5±4.7	0.80

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.

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.

АЛГОРИТМ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЯ ИЗОБРАЖЕНИЙ ЧЕЛОВЕЧЕСКИХ ЛИЦ, ОСНОВАННЫЕ НА МЕТОДЕ КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ В СИСТЕМЕ РАСПОЗНОВАНИЯ

Турапов У.У.
Кандидат технических наук, доцент. Ташкентский государственный аграрный университет, Ташкент, Республика Узбекистан
tuu_1224@gmail.com
Акилбаев М.И.
Кандидат наук, профессор. Университет дружбы народов, академик А. Каутбекова, Республика Казахстан
musabek_kz@mail.ru
Мадалиева Гульнар Уразалиевна
Старший преподаватель Университета дружбы народов имени академика А. Каутбековой, Республика Казахстан
Gulnar121974@mail.ru
Джантеева Шолпан Калмахановна
Старший преподаватель
Университета дружбы народов, академик А. Каутбекова, Республика Казахстан
shona2108@mail.ru

Аннотация: В настоящее время в данной статье представлены различные методы и комплекс прикладных программ для решения задач идентификации потоков видеоизображений, которые используются для решения задач биометрических систем. Авторы предложили создать автоматизированную систему анализатора глаз робота и использовать метод коэффициентной корреляции для идентификации изображения лица в полученных потоках цветных видеоизображений, и представлена функциональная схема этого метода.

Ключевые слова: цифровое изображение, технология Open MP, интенсивность, пиксель, векторизация, экстремальные точки, изображение лица.