

Метод объёмной сегментации и количественной оценки биомаркеров заболеваний сетчатки по последовательностям двумерных оптических когерентных томограмм

Е. Н. Волков

ФИЦ ИУ РАН

envolkoff@gmail.com

Аннотация. В докладе представлен метод объёмной сегментации и количественной оценки биомаркеров заболеваний сетчатки по последовательностям двумерных сканов оптической когерентной томографии. Подход основан на сегментации экземпляров на отдельных двухмерных срезах, сопоставлении объектов одного класса на смежных срезах и формировании пространственно-согласованных трёхмерных областей интереса с последующим вычислением их количественных характеристик. В качестве нейросетевой модели используется модифицированная архитектура YOLOv8m-seg-BiFPN-CA с усиленным межмасштабным слиянием признаков, сохранением их пространственной локализации. Предложенный конвейер ориентирован на работу в условиях ограниченной трёхмерной разметки и позволяет перейти от набора независимых посрезовых масок к целостному описанию патологических структур. Показана применимость метода для анализа биомаркеров диабетического макулярного отёка и расчёта их объёма на независимых данных внешней валидации.

Ключевые слова: искусственный интеллект; искусственные нейронные сети; нейросети; объёмная сегментация; сегментация экземпляров; трёхмерные медицинские изображения; оптическая когерентная томография

I. ВВЕДЕНИЕ

Последовательности двумерных сканов оптической когерентной томографии (ОКТ, ОК-томограммы) содержат биомаркеры, значимые для выявления патологических изменений, оценки активности процесса и контроля ответа на лечение при различных заболеваниях сетчатки глаза [4]. Наиболее полное описание таких структур даёт их трёхмерное представление, поскольку оно позволяет учитывать пространственную протяжённость очагов и переходить от локализации к количественной оценке объёма [5]. Однако прямой переход к полноценной трёхмерной сегментации обычно связан не только с использованием архитектур нейронных сетей, реализующих трёхмерную сегментацию (3D U-Net, V-Net, nnU-Net и др.), но и с высокой зависимостью от детально размеченных данных и вычислительных ресурсов, что в реальной практике далеко не всегда реализуемо [1-3]. Поэтому анализ двумерных срезов сохраняет практическую значимость. Вместе с тем, изолированная сегментация отдельных срезов не обеспечивает согласованного восстановления пространственной структуры и ограничивает корректную количественную интерпретацию областей интереса [5]. В настоящем докладе предлагается метод,

основанный на двумерной сегментации экземпляров, сопоставлении смежных срезов и формировании количественно описанных трёхмерных областей интереса на последовательностях двумерных сканов оптической когерентной томографии

II. МЕТОДЫ ИССЛЕДОВАНИЯ

A. Набор данных

В работе используется ранее сформированный собственный набор данных двухмерных срезов ОКТ для задачи сегментации экземпляров биомаркеров диабетического макулярного отёка. Набор включает 900 ОК-томограмм разрешением 512×1024 пикселя и содержит 6752 экземпляра биомаркеров восьми классов: HE, HF, IRC, EZD, SRF, VMT, DRIL, ERM. Среднее число аннотаций на изображение составляет 9,6. Формирование набора выполнено полуавтоматическим методом аннотирования, который объединяет предварительную разметку базовой моделью, автоматизированную оценку качества изображений и двухэтапную экспертную верификацию, завершающуюся построением согласованной клинически верифицированной разметки [6].

B. Архитектура YOLOv8m-seg-BiFPN-CA

В качестве базовой модели сегментации экземпляров используется YOLOv8m-seg, в которой иерархия многоуровневого слияния признаков дополнена модулем двунаправленной пирамиды признаков (BiFPN) [7], а в блоки обработки признаков встроен модуль Coordinate Attention (CA) [8]. Комбинация этих модулей обеспечивает более эффективное межмасштабное слияние признаков с сохранением информации о пространственном распределении информативных активаций, что особенно важно при выделении малых по площади и слабоконтрастных биомаркеров.

В BiFPN слияние карт признаков описывается как:

$$w_i^* = \frac{\max(0, w_i)}{\varepsilon + \sum_{j=1}^M \max(0, w_j)}$$
$$Y = \sum_{i=1}^M w_i^* X_i.$$

В модуле Coordinate Attention для входного тензора X вычисляются два направленных одномерных описания признаков, получаемых усреднением по горизонтальным и по вертикальным координатам:

$$z_c^h(y) = \frac{1}{W} \sum_{x=1}^W X_c(y, x),$$

$$z_c^w(x) = \frac{1}{H} \sum_{y=1}^H X_c(y, x),$$

в последующем выходной тензор определяется как:

$$Y_c(y, x) = X_c(y, x) \cdot g_c^h(y) \cdot g_c^w(x).$$

Включение модуля Coordinate Attention улучшает передачу информации о тонких структурах между уровнями представления и способствует более точному выделению малых по площади биомаркеров.

Обучение модели выполняется на основе оптимизации комплексной функции потерь, заданной взвешенной суммой частных функций потерь:

$$L_{total} = \lambda_{cls} L_{cls} + \lambda_{box} L_{box} + \lambda_{mask} L_{mask},$$

где:

$$L_{cls} = L_{Focal},$$

$$L_{bbox} = L_{DIOU},$$

$$L_{mask} = L_{Focal-mask} + L_{Tversky}.$$

Фокальная функция потерь [9] задаётся в виде:

$$L_{Focal} = -\alpha (1 - p_t)^\gamma \log(p_t).$$

Для ограничивающих прямоугольников используется функция потерь на основе метрики IoU [10]:

$$L_{DIOU} = 1 - IoU(B, B_{pred}) + \frac{\rho^2(b, b_{pred})}{c^2}.$$

Для масок используется функция потерь Тверского [11], заданная в виде:

$$L_{Tversky} = 1 - \frac{TP + \varepsilon}{TP + \alpha_t FP + \beta_t FN + \varepsilon}.$$

Выбранная композиция функции потерь согласована с постановкой задачи, а именно фокальная составляющая уменьшает вклад легко классифицируемых примеров, составляющая DIOU стабилизирует регрессию ограничивающих прямоугольников при малом перекрытии, а функция потерь Тверского позволяет увеличить штраф за пропуск малых объектов.

C. Описание шагов предлагаемого метода

Шаг 1. Чтение файлов изображений. Пусть $F = \{f_q\}_{q=1}^Q$ есть множество файлов в формате DICOM. Оператор чтения $D(f_q) = (I_q, \mu_q)$ ставит в соответствие каждому файлу матрицу интенсивностей $I_q \in \mathbb{R}^{H_q \times W_q}$ и набор метаданных μ_q .

Шаг 2. Группировка по идентификатору исследования и идентификатору серии. Для каждой пары значений, определяющих исследование и серию, формируется подмножество вида:

$$S = (I_q, \mu_q) \mid q \in 1, \dots, Q, \alpha(\mu_q) = a, \beta(\mu_q) = b,$$

где $\alpha(\mu_q)$ – функция извлечения идентификатора исследования, $\beta(\mu_q)$ – функция извлечения идентификатора серии. Далее S рассматривается как набор срезов одного объёмного обследования.

Шаг 3. Сортировка по положению среза через метаданные положения и ориентации. Серия снимков допускает представление в виде упорядоченной последовательности с согласованной ориентацией. Для каждого среза извлекаются вектор положения $p_k \in \mathbb{R}^3$ и два направляющих вектора $u_k, v_k \in \mathbb{R}^3$, задающих ориентацию плоскости среза. Нормаль к плоскости рассчитывается как:

$$n_k = u_k \times v_k.$$

Вводится опорная нормаль n , после чего координата среза вдоль последовательности определяется как:

$$t_k = \langle p_k, n \rangle.$$

Срезы упорядочиваются по возрастанию t_k и переобозначаются как $\{(I_k, \mu_k)\}_{k=1}^T$.

Шаг 4. Извлечение пространственного шага по плоскости изображения и шага между срезами. Для каждого среза извлекаются размеры пикселя $\Delta_r^k > 0$ и $\Delta_c^k > 0$. Переход от индексов пикселей к координатам номинального пространства серии задаётся отображением вида:

$$\Phi_k(i, j) = p_k + i\Delta_r^k u_k + j\Delta_c^k v_k.$$

Шаг между срезами определяется как:

$$\Delta_z^k = |t_{k+1} - t_k|, \quad k = 1, \dots, T - 1.$$

В дальнейшем величины Φ_k , Δ_r^k , Δ_c^k и Δ_z^k трактуются как параметры номинального координатного пространства серии двухмерных снимков.

Шаг 5. Предподготовка изображения. Для каждого среза вводится оператор преобразования:

$$x_k = T(I_k), \quad x_k \in \mathbb{R}^{h \times w \times d},$$

где T переводит исходное изображение в представление, на основе морфологических преобразований.

Шаг 6. Сегментация каждого среза. Обучающая выборка для многоклассовой сегментации экземпляров задаётся как:

$$D = \{(x_n, y_n)\}_{n=1}^N, \quad y_n = \{(m_{nr}, c_{nr})\}_{r=1}^{K(x_n)},$$

где $m_{nr} \in \{0, 1\}^{h \times w}$ – бинарная маска r -го экземпляра на n -м изображении, $c_{nr} \in 1, \dots, C$ его класс, а $K(x_n)$ – число экземпляров на изображении.

Параметры отображения сегментации f_θ определяются как решение задачи оптимизации:

$$\hat{\theta} = \operatorname{argmin}_\theta \frac{1}{N} \sum_{n=1}^N L(f_\theta(x_n), y_n) + \lambda \Omega(\theta),$$

где L – функция потерь, инвариантная к перестановке порядка экземпляров, $\Omega(\theta)$ – регуляризирующий член целевой функции. После обучения для n -го среза получаем предсказание:

$$\hat{y}_k = f_{\hat{\theta}}(x_k) = \{(\hat{m}_{k,r}, \hat{c}_{k,r})\}_{r=1}^{\hat{K}_k},$$

где $\hat{m}_{k,r} \in \{0, 1\}^{h \times w}$, $\hat{c}_{k,r} \in 1, \dots, C$, \hat{K}_k число предсказанных экземпляров на срезе.

Шаг 7. Обратное отображение масок в исходную матрицу изображения. Поскольку дальнейшее построение выполняется по обычным изображениям,

вводится оператор обратного геометрического преобразования:

$$B_k: \{0,1\}^{h \times w} \rightarrow \{0,1\}^{H_k \times W_k}.$$

Тогда предсказания в координатах исходного изображения среза имеют следующий вид:

$$\bar{y}_k = \{(\bar{m}_{k,r}, \hat{c}_{k,r})\}_{r=1}^{\hat{K}_k}, \quad \bar{m}_{k,r} = B_k(\hat{m}_{k,r}),$$

где $\bar{m}_{k,r}$ – маска на исходной матрице изображения I_k .

Шаг 8. Сопоставление экземпляров между смежными срезами в номинальных координатах серии. Для соседних срезов k и $k+1$ вводится отношение соответствия:

$$R_k \subset 1, \dots, \hat{K}_k \times 1, \dots, \hat{K}_{k+1}.$$

Пара индексов (r, s) принадлежит R_k , если соответствующие экземпляры относятся к одному классу и согласованы в номинальном координатном пространстве серии, то есть $(r, s) \in R_k$, тогда и только тогда, когда $\hat{c}_{k,r} = \hat{c}_{k+1,s}$ и

$$d(\Phi_k(A_{k,r}), \Phi_{k+1}(A_{k+1,s})) \leq \tau,$$

где $A_{k,r}$ – область r -го экземпляра на k -м срезе, d – мера пространственного рассогласования на подмножествах \mathbb{R}^3 , $\tau > 0$ – порог согласованности.

Шаг 9. Формирование трёхмерных объектов. На множестве всех предсказанных экземпляров серии строится граф вида:

$$G = (V, E), \quad V = \{(k, r) \mid 1 \leq k \leq T, 1 \leq r \leq \hat{K}_k\}, \\ E = \{((k, r), (k+1, s)) \mid (r, s) \in R_k\}.$$

Максимальные связные компоненты графа $\{O_g\}_{g=1}^M$ определяют трёхмерные объекты. Поскольку связи допускаются только между экземплярами одного класса, каждый объект O_g имеет единственный класс c_g .

Шаг 10. Подсчёт объёмов в кубических миллиметрах. Пусть $m_k^{(g)}$ – объединённая бинарная маска объекта O_g на k -м срезе, а при отсутствии объекта на срезе $m_k^{(g)} = 0$. Тогда номинальная площадь объекта интереса (биомаркера) на срезе равна:

$$A_k^{(g)} = \sum_{i=1}^{H_k} \sum_{j=1}^{W_k} m_k^{(g)}(i, j) \Delta_r^k \Delta_c^k,$$

а его номинальный объём вычисляется на основе послышной дискретной аппроксимации:

$$V(O_g) = \sum_{k=1}^{T-1} \frac{A_k^{(g)} + A_{k+1}^{(g)}}{2} \Delta_z^k.$$

Если Δ_r^k , Δ_c^k и Δ_z^k выражены в миллиметрах, то $V(O_g)$ выражается в кубических миллиметрах в номинальном координатном пространстве серии.

Шаг 11. Сохранение результатов. Для сохранения результатов вводится общая дискретная сетка $\Gamma = 1, \dots, H \times 1, \dots, W$ и операторы приведения $S_k: \{0,1\}^{H_k \times W_k} \rightarrow \{0,1\}^{H \times W}$, $k = 1, \dots, T$, которые переводят маски всех срезов в единую сетку серии. После этого помеченный трёхмерный массив задаётся

как $L(i, j, k) = g$, если пиксель $(i, j) \in \Gamma$ на k -м срезе принадлежит объекту O_g , и $L(i, j, k) = 0$ в противном случае. Кроме того, формируется таблица $T = \{(g, c_g, V(O_g), M_g)\}_{g=1}^M$, где g – идентификатор трёхмерного объекта, c_g – класс, $V(O_g)$ – объём, M_g – вектор дополнительных количественных характеристик.

III. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для внешней валидации предложенного метода использованы открытые наборы данных, представленные в работах [12] и RETOUCH [16]. Указанные ресурсы представляют практический интерес для оценки переносимости алгоритма на независимые данные и для анализа качества сегментации жидкостных биомаркеров в объёмном представлении. С учётом постановки настоящей работы основное внимание уделено биомаркерам IRC, IRF и SRF, поскольку именно они наиболее показательны для оценки точности сопоставления смежных срезов, пространственной согласованности реконструкции и корректности последующего расчёта объёма.

Количественные показатели качества сегментации и количественной оценки объёма приведены в табл. 1. Наиболее высокие значения достигнуты для IRF, что указывает на высокую устойчивость метода при восстановлении внутрисетчаточных жидкостных полостей. Для IRC получены близкие показатели, что подтверждает применимость предложенного конвейера к сегментации жидкостных включений данного типа. Некоторое снижение точности для SRF представляется ожидаемым, поскольку субретинальная жидкость характеризуется большей вариабельностью формы и менее однозначной конфигурацией границ в последовательности срезов.

ТАБЛИЦА I. Показатели качества сегментации и количественной оценки объёма при внешней валидации (СРЕДНЕЕ ЗНАЧЕНИЕ \pm СТАНДАРТНОЕ ОТКЛОНЕНИЕ)

Биомаркер	Dice, %	IoU, %	Относительная абсолютная ошибка объёма, %
IRC	90,48 \pm 2,31	82,70 \pm 3,62	8,64 \pm 3,12
IRF	91,12 \pm 1,94	83,89 \pm 3,08	7,91 \pm 2,74
SRF	88,37 \pm 2,76	79,26 \pm 4,18	10,83 \pm 3,95

Сопоставление коэффициента Дайса (Dice) и индекса Жаккара (IoU) демонстрирует согласованную картину качества по всем трём биомаркерам. Во всех случаях достигнут уровень, соответствующий точному восстановлению формы патологической области в объёме, что позволяет рассматривать предложенный подход не только как средство сегментации отдельных срезов, но и как инструмент формирования пространственно согласованного трёхмерного представления области интереса. Дополнительную информацию даёт относительная абсолютная ошибка объёма, вычисляемая как отношение модуля разности между предсказанным и эталонным объёмом к эталонному объёму в процентах на основе формулы:

$$E_V = \frac{\text{abs}(V_{\text{pred}} - V_{\text{true}})}{V_{\text{true}}} \times 100\%,$$

где V_{pred} и V_{true} обозначают предсказанный и эталонный объёмы соответственно. Для IRF и IRC её средние

значения остаются ниже 10%, что указывает на хорошую согласованность сегментации и итогового количественного расчёта. Для SRF ошибка объёма несколько выше, что соответствует меньшему перекрытию масок и отражает более высокую чувствительность данного класса к погрешностям восстановления границ.

В совокупности полученные результаты подтверждают, что конвейер, основанный на сегментации экземпляров на двумерных срезах, сопоставлении смежных срезов и агрегации в трёхмерные объекты, обеспечивает высокое качество выделения жидкостных биомаркеров и приемлемую точность расчёта их объёма. Практическая значимость такого подхода состоит в переходе от набора независимых посрезовых масок к пространственно согласованному трёхмерному описанию патологического объекта, допускающему вычисление интегральных количественных характеристик и последующий количественный анализ динамики патологического процесса.

Дополнительным средством верификации качества обучения нейросетевой модели должен выступать объяснительный искусственный интеллект [13, 14]. Высокие значения интегральных метрик не гарантируют, что нейросетевая модель опирается именно на клинически значимые признаки. Анализ карт значимости позволяет проверить, локализуется ли вклад решения в пределах патологически релевантных областей, а не на вторичных текстурных или аппаратно-обусловленных особенностях изображения. Для трёхмерных архитектур применим метод 3D Grad-CAM, позволяющий локализовать участки объёмного представления, вносящие наибольший вклад в формирование предсказания [15]. Включение такого этапа в процедуру анализа делает интерпретацию результатов более обоснованной как с методической, так и с прикладной точек зрения.

IV. ЗАКЛЮЧЕНИЕ

В работе предложен метод объёмной сегментации и количественной оценки биомаркеров заболеваний сетчатки по последовательностям двумерных ОКТ-срезов, основанный на сегментации экземпляров, сопоставлении срезов и формировании пространственно-согласованных трёхмерных объектов. Результаты внешней валидации подтверждают, что предложенный конвейер обеспечивает высокое качество выделения жидкостных биомаркеров IRC, IRF и SRF и сохраняет приемлемую точность расчёта их объёма. Тем самым показано, что использование последовательности двумерных срезов может служить содержательно обоснованной основой для восстановления трёхмерной структуры патологических изменений и последующего морфометрического анализа, значимого для мониторинга течения заболевания и оценки эффективности лечения при невозможности использования классической трёхмерной сегментации.

СПИСОК ЛИТЕРАТУРЫ

- [1] Cicek O., Abdulkadir A., Lienkamp S. S., Brox T., Ronneberger O. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation // *Medical Image Computing and Computer-Assisted Intervention*. MICCAI 2016. Lecture Notes in Computer Science. 2016. Т. 9901. С. 424-432. DOI: 10.1007/978-3-319-46723-8_49.
- [2] Wang S., Li C., Wang R., Liu Z., Wang M., Tan H., Wu Y., Liu X., Sun H., Yang R., Liu X., Chen J., Zhou H., Ben Ayed I., Zheng H. Annotation-efficient deep learning for automatic medical image segmentation // *Nature Communications*. 2021. Т. 12. № 1. Ст. 5915. DOI: 10.1038/s41467-021-26216-9.
- [3] Zhang Y., Liao Q., Ding L., Zhang J. Bridging 2D and 3D segmentation networks for computation-efficient volumetric medical image segmentation: An empirical study of 2.5D solutions // *Computerized Medical Imaging and Graphics*. 2022. Т. 99. Ст. 102088. DOI: 10.1016/j.compmedimag.2022.102088.
- [4] Pandya B. U., Grinton M., Mandelcorn E. D., Felfeli T. Retinal optical coherence tomography imaging biomarkers: a review of the literature // *Retina*. 2024. Т. 44. № 3. С. 369-380. DOI: 10.1097/IAE.0000000000003974.
- [5] Guo Y., Hormel T. T., Xiong H., Wang J., Hwang T. S., Jia Y. Automated Segmentation of Retinal Fluid Volumes From Structural and Angiographic Optical Coherence Tomography Using Deep Learning // *Translational Vision Science & Technology*. 2020. Т. 9. № 2. Ст. 54. DOI: 10.1167/tvst.9.2.54.
- [6] Volkov E. Semi-Automated Method for Annotating Biomedical Images Using Artificial Intelligence Technologies // 2025 VI International Conference on Control in Technical Systems (CTS). IEEE, 2025. С. 135-138. DOI: 10.1109/CTS67336.2025.11196701.
- [7] Tan M., Pang R., Le Q. V. EfficientDet: Scalable and Efficient Object Detection // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020. P. 10781-10790. DOI: 10.1109/CVPR42600.2020.01079.
- [8] Hou Q., Zhou D., Feng J. Coordinate Attention for Efficient Mobile Network Design // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021. P. 13713-13722. DOI: 10.1109/CVPR46437.2021.01350.
- [9] Lin T.-Y., Goyal P., Girshick R., He K., Dollár P. Focal Loss for Dense Object Detection // 2017 IEEE International Conference on Computer Vision. 2017. P. 2999-3007. DOI: 10.1109/ICCV.2017.324.
- [10] Zheng Z., Wang P., Liu W., Li J., Ye R., Ren D. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2020. Vol. 34, No. 7. P. 12993-13000. DOI: 10.1609/aaai.v34i07.6999.
- [11] Salehi S. S. M., Erdogmus D., Gholipour A. Tversky Loss Function for Image Segmentation Using 3D Fully Convolutional Deep Networks // *Machine Learning in Medical Imaging*. 2017. P. 379-387. DOI: 10.1007/978-3-319-67389-9_44.
- [12] Huang W., Qin L., Xu M., Zheng H., Gan Y., Pei S., Wu R., Liu Y., Zhong J., Ni G. Comprehensive 3D Optical Coherence Tomography Dataset for AMD and DME: Facilitating Deep-Learning-Based 3D Segmentation // *Scientific Data*. 2026. Т. 13. Ст. 224. DOI: 10.1038/s41597-025-06497-1.
- [13] Averkin A.N., Volkov E.N., Yarushev S.A. Explainable artificial intelligence in deep learning neural nets-based digital images analysis // *Journal of Computer and Systems Sciences International*. 2024. Т. 63. № 1. С. 175-203. DOI: 10.1134/S1064230724700138.
- [14] Volkov E.N., Averkin A.N. Explainable artificial intelligence in medical image analysis: State of the art and prospects // 2023 XXVI International Conference on Soft Computing and Measurements (SCM). IEEE, 2023. С. 134-137. DOI: 10.1109/SCM58628.2023.10159033.
- [15] Huo Y., Terry J. G., Wang J., Nath V., Bermudez C., Bao S., Parvathaneni P., Carr J. J., Landman B. A. Coronary Calcium Detection using 3D Attention Identical Dual Deep Network Based on Weakly Supervised Learning // *Proceedings of SPIE*. 2019. Т. 10949. Ст. 1094917. DOI: 10.1117/12.2512541.
- [16] Bogunovic H., Venhuizen F., Klmscha S., Apostolopoulos S., Babadiashar A., Bagci U., Beg M. F., Bekalo L., Chen Q., Ciller C., Gopinath K., Gostar A. K., Jeon K., Ji Z., Kang S. H., Koozekanani D. D., Lu D., Morley D., Parhi K. K., Park H. S., Rashno A., Sarunic M., Shaikh S., Sivaswamy J., Tennakoon R., Yadav S., De Zanet S., Waldstein S. M., Gerendas B. S., Klaver C., Sanchez C. I., Schmidt-Erfurth U. RETOUCH: The Retinal OCT Fluid Detection and Segmentation Benchmark and Challenge // *IEEE Transactions on Medical Imaging*. 2019. Т. 38. № 8. С. 1858-1874. DOI: 10.1109/TMI.2019.2901398