

Модификация архитектуры U-Net на основе Габоровых фильтров с обучаемыми параметрами для улучшения сегментации биомаркеров на ОКТ-снимках

Е. Н. Волков, А. Н. Аверкин

Государственный университет «Дубна»

envolkoff@gmail.com

Аннотация. Оптическая когерентная томография (ОКТ) широко используется для диагностики патологий сетчатки, однако низкий контраст и зашумленность ОКТ-изображений часто затрудняют автоматическую сегментацию биомаркеров. В данной работе ставится цель повысить точность автоматической сегментации биомаркеров ОКТ путём модификации архитектуры U-Net. Предложена архитектура Gabor-U-Net на основе U-Net, интегрирующая обучаемые Габоровы фильтры в свёрточные слои для улучшения распознавания ориентированных текстур и границ. Эксперименты проведены на датасете OIMHS. Качество сегментации оценивалось метриками IoU и коэффициентом Dice. Результаты показывают, что Gabor-U-Net достоверно превосходит классическую U-Net по значениям метрик: для сегментации МН достигнуты значения IoU=0.90 и Dice=0.94; для IRC – IoU=0.88 и Dice=0.91. Это на 3–4 % выше показателей базовой модели, что подтверждает эффективность предложенной модификации для точной сегментации диагностически важных биомаркеров на ОКТ-изображениях.

Ключевые слова: искусственный интеллект, нейросети, сегментация, биомаркеры, окт, офтальмология, объяснительный искусственный интеллект

I. ВВЕДЕНИЕ

Оптическая когерентная томография (ОКТ) широко применяется в офтальмологии для неинвазивного исследования сетчатки. Сегментация биомаркеров на ОКТ-снимках (например, отёков сетчатки, отложений или границ слоёв) имеет высокую клиническую и диагностическую значимость, так как обеспечивает объективное количественное измерение патологических изменений и динамики заболевания [1]. Особенно при таких патологиях как возрастная макулярная дегенерация точное выделение структур позволяет отслеживать прогрессирование болезни и оценивать эффективность терапии. Однако ручная сегментация ОКТ-изображений трудоёмка и субъективна, поэтому автоматизация сегментации с помощью методов глубокого обучения привлекает всё больше внимания исследователей [2].

Архитектура U-Net стала одним из базовых подходов для сегментации медицинских изображений благодаря своей U-образной энкодер-декодер структуре с

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № 124112200072-2, Применение объяснительного искусственного интеллекта для интерпретации алгоритмов машинного обучения)

пропускными связями, эффективно объединяющей контекст и детали. Многочисленные работы демонстрируют успешное применение U-Net и её модификаций для сегментации ретинальных слоёв и патологических образований на ОКТ. Тем не менее, стандартные свёрточные сети могут испытывать затруднения при распознавании тонких текстурных особенностей и ориентированных структур в изображениях. Это мотивирует внедрение в модель специализированных фильтров, чувствительных к ориентации границ и текстур.

Габоровы фильтры давно известны способностью эффективно выделять текстурные признаки и ориентированные градиенты изображения, имитируя работу рецептивных полей зрительной коры. Интеграция таких фильтров в архитектуру U-Net с возможностью обучения их параметров (пространственной частоты, масштаба, ориентации и фазового сдвига) позволяет модели адаптивно подстраиваться под характерные структуры данных [3]. Предыдущие исследования показали, что использование обучаемых Габоровых фильтров в слоях CNN улучшает качество сегментации вытянутых и направленных объектов, таких как кровеносные сосуды сетчатки [4]. Ожидается, что предлагаемая модификация U-Net, дополняющая стандартные свёрточные слои обучаемыми Габоровыми фильтрами, усилит выделение тонких ориентированных особенностей и текстур, благодаря чему повысится точность сегментации диагностически важных биомаркеров на ОКТ-снимках.

II. МЕТОДЫ

A. Фильтры Габора

Фильтр Габора представляет собой линейный фильтр, задаваемый как комплексная гармоническая функция, локализованная гауссовым окном [5]. Его двумерная форма обычно записывается в виде (1):

$$g(x, y, \omega, \theta, \psi, \sigma) = \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \exp(i(\omega x' + \psi)), \quad (1)$$

$$x' = x \cos \theta + y \sin \theta, \quad (2)$$

$$y' = -x \sin \theta + y \cos \theta, \quad (3)$$

где ω – угловая частота синусоиды, θ – угол ориентации фильтра, ψ – фазовый сдвиг (начальная фаза), σ – диапазон (стандартное отклонение гауссианы).

Вещественная часть этого фильтра получается заменой экспоненты на косинус [7] (4):

$$g(x, y, \omega, \theta, \psi, \sigma) = \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \cos(\omega x' + \psi). \quad (4)$$

Благодаря такой форме, Габоровы фильтры выполняют локализованный частотно-пространственный анализ изображения и эффективно выделяют текстурные признаки и ориентированные градиенты [6]. Такое свойство делает их особо подходящими для задач сегментации текстур, например, на ОКТ-снимках, где диагностические биомаркеры проявляются как тонкие слоистые структуры или неоднородности, различающиеся по частоте и направлению. Известно, что первые слои обученных СНС спонтанно формируют фильтры, близкие по форме к Габоровым, что свидетельствует о важности таких признаков. Поэтому имеет смысл явно включить Габоровы фильтры в архитектуру нейросети [8].

В. Архитектура U-Net

Архитектура U-Net была предложена Ronneberger О. в 2015 году и представляет собой симметричную энкодер-декодер сеть, специально разработанную для задач семантической сегментации медицинских изображений и характеризующуюся U-образной топологией с пропускными соединениями (skip-connections) между соответствующими уровнями кодировщика и декодировщика [9] (рис. 1).

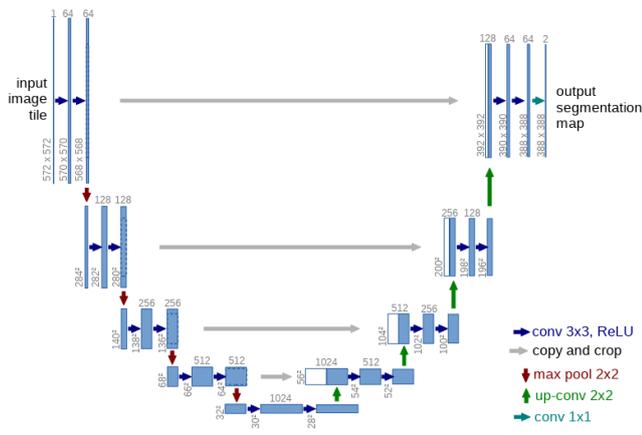


Рис. 1. Архитектура U-Net [9]

На контрактурирующем (энкодирующем) пути последовательно выполняются операции свёртки и субдискретизации, что позволяет сети извлекать контекстные признаки различных масштабов. Формально выход сверточного слоя на уровне l можно записать как (5):

$$F^{(l)} = \sigma(W^{(l)} * F^{(l-1)} + b^{(l)}), \quad (5)$$

где $F^{(l-1)}$ – входные карты признаков, $W^{(l)}$ и $b^{(l)}$ – ядро свёртки и смещение, символ «*» – операция свёртки, а σ – функция активации (ReLU). Субдискретизация реализуется через операцию max-ooling с окном 2x2, уменьшающую пространственные размеры вдвое и позволяющую захватить больше глобальных контекстов.

На декодирующем пути архитектуры выполняется апсемплинг посредством транспонированных свёрток, формирующих $\tilde{F}^{(l)}$ – карту признаков восстановленного разрешения (в два раза выше по сравнению с предыдущим уровнем). К каждой $\tilde{F}^{(l)}$ присоединяются (concat) соответствующие по разрешению карты $F^{(l)}$ из энкодера, обеспечивая передачу высокочастотных деталей, утраченных при пулинге. Таким образом, для каждого уровня декодирования выполняется (6, 7):

$$U^{(l)} = \sigma(W_{up}^{(l)} * \tilde{F}^{(l)} + b_{up}^{(l)}), \quad (6)$$

$$D^{(l)} = \sigma(W_{dec}^{(l)} * [U^{(l)}, F^{(l)}] + b_{dec}^{(l)}), \quad (7)$$

где $[\cdot; \cdot]$ – конкатенация по каналному измерению, а $W_{up}^{(l)}, W_{dec}^{(l)}, b_{up}^{(l)}, b_{dec}^{(l)}$ – обучаемые параметры соответствующих слоёв декодера. Заключительный сверточный слой проецирует полученные карты признаков D в пространство меток сегментации, выдавая на выход карту вероятностей принадлежности каждого пикселя к заданным классам.

Преимущество U-Net в задачах медицинской сегментации заключается в эффективном объединении локальной детализации и глобального контекста при сравнительно небольшом числе обучающих образцов. Пропускные соединения помогают сохранить тонкие структурные признаки (границы, текстуры) и обеспечивают более быстрое и стабильное распространение градиента во время обучения, что критично при глубокой иерархии слоёв. Эксперименты показали, что U-Net достигает высокой точности сегментации кровеносных сосудов, опухолевых масс и других биомаркеров в ОКТ- и МРТ-изображениях, часто превосходя классические методы, такие как FCN и SegNet, по метрике Dice и IoU при одинаковом объёме данных и вычислительных ресурсов. [10]

Кроме того, U-Net демонстрирует высокую гибкость: архитектура легко расширяется за счёт добавления глубины, ширины или внедрения современных модулей (attention, резидуальные блоки, трансформеры), сохраняя исходную идею симметричного строения с пропускными связями. Благодаря этим свойствам U-Net стала стандартом для медицинской сегментации и активно применяется в клинических и исследовательских задачах, требующих точного выделения биомаркеров на изображениях с низким контрастом и высоким уровнем шума. [11]

С. Набор данных

Датасет OIMHS (Optical Coherence Tomography Image Dataset Based on Macular Hole Manual Segmentation) представляет собой высококачественную коллекцию из 3859 В-сканов ОКТ-изображений, полученных от 119 пациентов (125 глаз) с подтверждённым макулярным отверстием. Каждое изображение сопровождается четырьмя вручную аннотированными масками: ретина (Retina), макулярное отверстие (MH), внутриканевые кисты (IRC) и хориоидея (Choroid). Примеры изображений датасета OIMHS представлены на рис. 2.

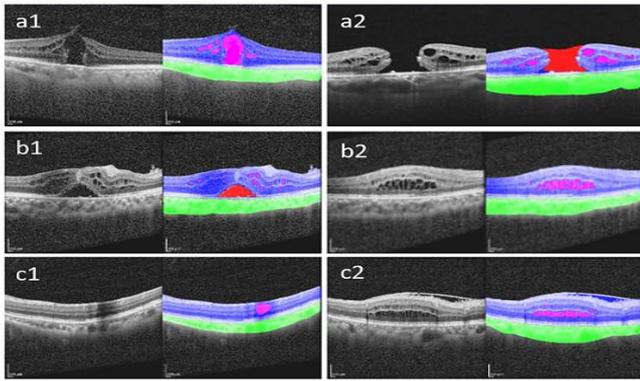


Рис. 2. Пример изображений и масок классов датасета OIMHS [12]

Первичная разметка выполнялась тремя младшими офтальмологами с использованием Clip Studio Paint, после чего экспертный офтальмолог проверял и корректировал все сегментации, присваивая каждой аннотации оценку качества (0–2 балла) по критериям точности и полноты. Дополнительно специалисты оценивали качество исходных OCT-кадров по трём параметрам: сила сигнала (низкая/высокая), наличие артефактов и размытость изображения, что обеспечивает учёт влияния качества сканов на устойчивость алгоритмов сегментации [12].

Для использования в данной работе были выбраны биомаркеры, макулярное отверстие (МН) и внутритканевые кисты (IRC).

D. Метрики оценки качества

В качестве метрик оценки качества в задаче семантической сегментации на несколько классов применялись Intersection over Union (8) и Dice Coeff. (9):

$$IoU_c = \frac{|P_c \cap G_c|}{|P_c \cup G_c|}, \quad (8)$$

$$Dice_c = \frac{|P_c \cap G_c|}{|P_c| + |G_c|}, \quad (9)$$

где P_c – множество пикселей, предсказанных как класс c ; G_c – множество пикселей истинной разметки (англ. ground truth) для класса c ; C – число классов.

III. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Анализ данных Таблицы I и кривых IoU (Рис. 3) и Dice (Рис. 4), построенных по эпохам обучения (50 эпох, средние значения $\pm 1\sigma$), демонстрирует устойчивую сходимость и стабильность обучения обеих моделей. К концу обучения значения метрик выходят на плато, а разброс показателей между запусками невелик, что свидетельствует о воспроизводимости результатов. На протяжении всего процесса обучения Gabor-U-Net показывает более высокие значения IoU и Dice по сравнению с классическим U-Net, указывая на его систематическое преимущество.

ТАБЛИЦА I. СРАВНЕНИЕ ЗНАЧЕНИЙ МЕТРИК

Нейросеть	Биомаркеры	IoU	Dice
U-Net (base)	IRC	0.84	0.89
	МН	0.86	0.92
Gabor-U-Net	IRC	0.88	0.91
	МН	0.90	0.94

Итоговые показатели метрик также подтверждают превосходство Gabor-U-Net над базовой архитектурой. Для биомаркера IRC средний IoU достиг 0.88 у Gabor-U-Net против 0.84 у U-Net, а для МН – 0.91 против 0.89. Аналогичная динамика наблюдается по Dice: значение увеличилось с 0.86 до 0.90 (IRC) и с 0.92 до 0.94 (МН) при переходе от U-Net к Gabor-U-Net. Разница особенно заметна для более сложного случая IRC, где выигрыш по метрикам выше. Эти результаты указывают на более высокую точность сегментации при использовании модели с Габоровыми фильтрами.

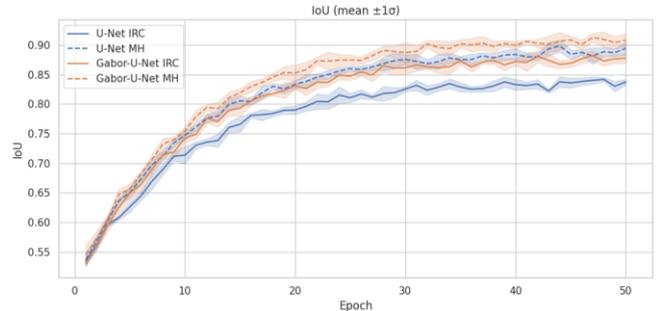


Рис. 1. Значения метрики IoU

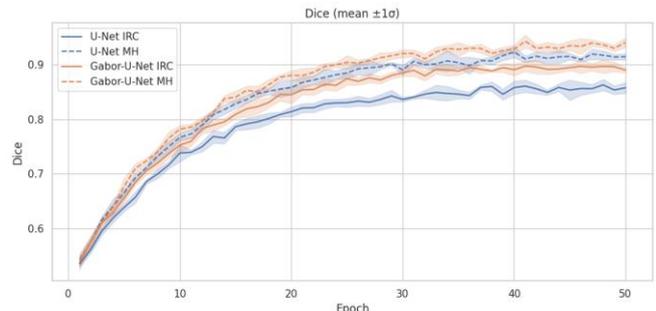


Рис. 2. Значения метрики Dice

Архитектурные особенности Gabor-U-Net, в частности внедрение обучаемых Габоровых фильтров, обеспечивают лучшее распознавание текстур и ориентированных градиентов в ОКТ-снимках. Благодаря этим фильтрам, уже на ранних сверточных слоях Gabor-U-Net эффективно выделяет локальные текстурные признаки биомаркеров, позволяя модели точнее отделять области IRC и МН от окружающих тканей. Анализ матрицы ошибок подтверждает, что Gabor-U-Net значительно снижает количество неправильных классификаций по сравнению с U-Net. В частности, новая модель реже допускает ложноположительные срабатывания (когда здоровые области ошибочно маркируются как патология) и пропускает меньше истинных поражений (ложноотрицательные ошибки). Снижение таких ошибок непосредственно отражается на росте показателей IoU и Dice.

В экспериментальном пайплайне сегментации ОКТ-изображений на основе Gabor-U-Net интеграция методов объяснительного искусственного интеллекта (XAI) существенно повышает интерпретируемость результатов модели. Основной подход заключается в анализе финальных предсказаний сети посредством визуализации тепловых карт, выделяющих ключевые области исходного изображения, наиболее влияющие на решение нейросети [13]. Применение алгоритмов XAI, например Grad-CAM, наглядно показывает, какие

признаки учитывает Gabor-U-Net при сегментации, тем самым обеспечивая прозрачность вывода модели. Такая поясняющая визуализация не только облегчает анализ сегментированных структур, но и помогает выявлять возможные ошибки или смещения модели, повышая доверие к её предсказаниям [14].

Таким образом, сочетание U-Net с чувствительными к текстурным признакам фильтрами Габора приводит к статистически значимому улучшению качества сегментации биомаркеров IRC и МН на ОКТ-снимках. На практике Gabor-U-Net обеспечивает более точное и устойчивое выделение патологических структур по сравнению со стандартной U-Net, что подтверждается более высокими значениями метрик IoU и Dice.

IV. ЗАКЛЮЧЕНИЕ

Проведённый сравнительный анализ классической U-Net и созданной модификацией – Gabor-U-Net продемонстрировал явное преимущество предлагаемого подхода. Интеграция обучаемых Габоровых фильтров позволяет Gabor-U-Net более эффективно выявлять ориентированные структуры и тонкие текстурные особенности на ОКТ-снимках, что приводит к систематическому повышению метрик сегментации (IoU, Dice) по сравнению с базовой моделью. Особенно заметен выигрыш Gabor-U-Net на более сложном случае – сегментации внутритканевых кист. Предложенная модификация обеспечивает более точное и воспроизводимое выделение диагностически значимых биомаркеров сетчатки, повышая объективность оценки патологических изменений. Полученные результаты имеют важное значение для развития автоматизированного анализа ОКТ: повышение качества сегментации способствует более надёжному мониторингу заболеваний и обоснованному принятию клинических решений. Перспективы дальнейших исследований включают интеграцию разработанного метода в клинический рабочий процесс анализа ОКТ-изображений для поддержки врачей-офтальмологов, а также адаптацию данного подхода к сегментации других патологий сетчатки (например, диабетического отёка макулы) и к другим модальностям медицинской визуализации.

СПИСОК ЛИТЕРАТУРЫ

- [1] Asani B., Holmberg O., Schiefelbein J.B. et al. Evaluation of OCT biomarker changes in treatment-naive neovascular AMD using a deep semantic segmentation algorithm // *Eye*. 2024. P. 1-7. DOI: 10.1038/s41433-024-03264-1.
- [2] Zhang H., Yang B., Li S. et al. Retinal OCT image segmentation with deep learning: A review of advances, datasets, and evaluation metrics // *Computerized Medical Imaging and Graphics*. 2025. P. 102539. DOI: 10.1016/j.compmedimag.2025.102539.
- [3] Alekseev A., Bobe A. GaborNet: Gabor filters with learnable parameters in deep convolutional neural network // 2019 International Conference on Engineering and Telecommunication (EnT). IEEE, 2019. P. 1-4. DOI: 10.1109/EnT47717.2019.9030571.
- [4] Yang C., Huang X., Tang L. et al. RAGE-Net: Enhanced retinal vessel segmentation U-shaped network using Gabor convolution // *Digital Signal Processing*. 2024. Vol. 153. P. 104643. DOI: 10.1016/j.dsp.2024.104643.
- [5] Chen C., Zhou K., Qi S. et al. A learnable Gabor Convolution kernel for vessel segmentation // *Computers in Biology and Medicine*. – 2023. Vol. 158. P. 106892. DOI: 10.1016/j.compbiomed.2023.106892.
- [6] Eghbali N., Bagher-Ebadian H., Alhanai T. et al. GLoG-CSUNet: Enhancing Vision Transformers with Adaptable Radiomic Features for Medical Image Segmentation // *arXiv preprint arXiv:2501.02788*. 2025.
- [7] Meshgini S., Aghagolzadeh A., Seyedarabi H. Face recognition using gabor filter bank, kernel principle component analysis and support vector machine // *International Journal of Computer Theory and Engineering*. 2012. Vol. 4. No. 5. P. 767. DOI: 10.7763/IJCTE.2012.V4.574.
- [8] Luan S., Chen C., Zhang B. et al. Gabor convolutional networks // *IEEE Transactions on Image Processing*. 2018. Vol. 27. No. 9. P. 4357-4366. DOI: 10.1109/TIP.2018.2835143.
- [9] Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // *Medical image computing and computer-assisted intervention – MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. – Springer international publishing, 2015. P. 234-241. DOI: 10.1007/978-3-319-24574-4_28.
- [10] Zhou Z., Siddiquee M.M.R., Tajbakhsh N. et al. Unet++: A nested unet architecture for medical image segmentation // *Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, DLMIA 2018, and 8th international workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, proceedings 4*. – Springer International Publishing, 2018. P. 3-11. DOI: 10.1007/978-3-030-00889-5_1.
- [11] Faes L., Lin H., Wagner S.K. et al. Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study // *The Lancet Digital Health*. 2019. Vol. 1. No. 5. P. e232-e242. DOI: 10.1016/S2589-7500(19)30108-6.
- [12] Ye X., He S., Zhong X. et al. Oimhs: An optical coherence tomography image dataset based on macular hole manual segmentation // *Scientific Data*. 2023. Vol. 10. No. 1. P. 769. DOI: 10.1038/s41597-023-02675-1.
- [13] Volkov E.N. and Averkin A.N., "Explainable Artificial Intelligence in Clinical Decision Support Systems," 2023 IV International Conference on Neural Networks and Neurotechnologies (NeuroNT), Saint Petersburg, Russian Federation, 2023, pp. 3-6, doi: 10.1109/NeuroNT58640.2023.10175852.
- [14] 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. Vol.63. No. 1. P. 175-203. DOI: 10.1134/S1064230724700138.