

# Сегментация опухолей головного мозга с использованием глубокого обучения

Нахида Каража

Санкт-Петербургский государственный  
электротехнический университет  
«ЛЭТИ» им. В.И. Ульянова (Ленина)

nahidahka@gmail.com

Яссер Низамли

Санкт-Петербургский государственный  
электротехнический университет  
«ЛЭТИ» им. В.И. Ульянова (Ленина)

yanizamli@stud.etu.ru

**Аннотация.** Сегментация опухолей, расположенных в полости головы – важнейший аспект медицинской визуализации, который позволяет врачам формировать точный диагноз, планировать лечение и предсказывать вероятность выздоровления пациента. Изготовление сегментаций вручную трудно, это требует времени, и зависит от специальности специалиста. Актуальные решения – глубокое обучение (DL) и методы его реализации, такие как свёрточные нейронные сети (CNN), различные варианты U-Net, трансформеры и их гибридные модели, наконец, позволяют получить автоматическую и точную маркировку областей опухоли на изображениях МРТ. Последние исследования показали, что DL-методы более точны и быстрее, чем ручная работа или классические методы машинного обучения. Например, лучшие модели достигают значений по Дайсу больше 0.9 на эталонах, таких как BraTS. Однако надо преодолевать возникающие сложности с обобщаемостью для разных протоколов визуализаций, интерпретируемостью, нехваткой данных и внедрением в клинический рабочий процесс. В данном обзоре собран материал последних работ, чтобы создать целостное представление о сегментации опухолей головного мозга на базе DL.

**Ключевые слова:** сегментация опухолей головного мозга, глубокое обучение, магнитно-резонансная томография (МРТ), свёрточные нейронные сети (CNN), U-Net, трансформеры, гибридные модели

## I. ВВЕДЕНИЕ

Головной мозг, опухоли в нем и опухоли других частей организма, представляющие собой нервную систему, занимают 10-е место среди самых распространённых причин смерти в развитых странах. Эта болезнь может поразить представителя любого возраста, как ребёнка, так и взрослого [1]. На МРТ опухоли обычно делятся на три части: само образование, ядро образования и области, ответственные за усиление изображения при контрастировании.

Сегментация опухолей головного мозга — это важная составляющая анализа медицинских изображений, которая позволяет врачам ставить точный диагноз, планировать дальнейшее лечение и прогнозировать вероятность благоприятного исхода для пациента. Ручная же сегментация требует много времени, и её результат зависит от конкретного человека. Поэтому для автоматического и точного определения областей опухоли на МРТ-сканах применяются методы глубокого обучения (DL). В частности, свёрточные нейронные сети (CNN) и их

производные U-Net, трансформеры и гибридные модели [2].

Недавние исследования показывают, что методы глубокого обучения более точны и быстры, чем традиционные методы машинного обучения и ручные методы. Так, современные модели показывают Dice больше 0.9 на референсных наборах BraTS [3]. Все же проблемы остаются, и они касаются обобщаемости результатов для различных протоколов визуализации, интерпретируемости даже самих моделей, нехватки данных и интеграции в клинические рабочие процессы. В данном обзоре собрано всё существующее о сегментации опухолей головного мозга с использованием методов DL [2][3].

## II. ОСНОВНЫЕ АРХИТЕКТУРЫ ГЛУБОКОГО ОБУЧЕНИЯ

В настоящее время в исследованиях по сегментации опухолей мозга на МРТ применяются классические CNN, варианты U-Net, трансформерные модели и гибридные подходы, степень совершенства которых постоянно увеличивается. Эти варианты U-Net, такие как 3D U-Net, Attention U-Net, а также модели, усовершенствованные за счёт базовых сетей типа VGG-16 или EfficientNet, и по-прежнему хорошо востребованы, поскольку они легко поддаются оптимизации и довольно успешно извлекают локальные признаки. Модели на основе трансформеров и гибридные модели опираются на механизмы самовнимания для нахождения дальних зависимостей, тем самым повышая точность сегментации, благодаря интеграции глобального контекста с локальными признаками. Такими моделями послужили TransBTS, Swin, UNETR, CA-U-Net с показателями выше, чем у обычного U-Net, но и потребляющие больше вычислительных мощностей. Последние исследования показали, что лучший способ найти оптимальное решение для задачи BraTS заключается в применении гибридных архитектур, увязывающих CNN-кодеры с трансформерными модулями. Причина этому — сбалансированность мелких пространственных деталей и глобального семантического понимания, что и даёт выигрыш. Эти модели также хорошо подходят для использования механизмов внимания и многомасштабного слияния признаков, что приводит к более точному распознаванию границ опухолей и границ подрегионов опухолей.

### A. Варианты U-Net и модели с улучшёнными CNN

- Оптимизированные версии семейства U-Net: Сравнительные эксперименты показывают, что

3D U-Net, Attention U-Net, R2-Attention U-Net и современные модификации U-Net показывают показатели Dice, сопоставимые с результатами, показанными лучшими моделями. Благодаря достаточно тщательной оптимизации на данном уровне даже простая U-Net по качеству работы превосходит гораздо более сложные архитектуры [4][5].

- U-Net с усовершенствованной базовой сетью: кодировщик U-Net + VGG-16 лучше, чем подкручиваться к U-Net на TCGA-LGG, когда речь идет о пиксельной точности [6]. U-Net с многомасштабным вниманием + EfficientNetB4 достигает Dice 0,934 и IoU 0,88 на Figshare за счёт синергии многомасштабных CNN-особенностей и многомасштабного внимания с фокусом на границах опухоли [7]. ABI-Net использует 3D-Inception и механизм внимания в U-Net, чтобы улучшить показатель Dice для подрегионов на наборе BraTS 2020 [8].

И на данный момент, как показывают обзоры, версии U-Net остаются самыми популярными моделями сегментации МРТ в медицине, особенно для мозга [4][5].

### *В. Трансформеры и гибридные модели CNN-Transformer*

а) Чистый / с преобладанием трансформатора: EF-VPT-Net использует трансформаторный визуальный патч-энкодер и U-Net сегментатор на основе свёрточной сети внутри декодера. Он может достигать точности сегментации порядка ~99 % на двух наборах данных МРТ за счёт раннего учёта дальнедействующих зависимостей [9].

#### *б) Канонические комбинации:*

- TransBTS — это 3D-кодер CNN, "бутылочное горлышко" трансформера и CNN декодер, который улучшает производительность BraTS 2019/2020 за счёт объединения локального 3D-контекста с глобальной семантикой [10].
- Swin UNETR — это модель, занявшая первое место в BraTS 2021, имеющая кодер Swin-трансформера, CNN декодер и U-образные пропускные соединения [11].
- 3DUV-NetR+: 3D-кодеры U-Net и V-Net, а затем трансформерные блоки в 3D. Приносит оценку DSC ~ 0.92 (WT) на BraTS 2020 [12].
- CA-U-Net и AugTransU-Net добавляют модули трансформера и расширенного трансформера к U-Net, а также и средства улучшения контекста (atrous conv и переработанные пропуски). И тем не менее оказываются значительно лучше, чем исходный U-Net по показателям IoU / DSC на данных типа BraTS [13] [14].
- TuSegNet (CNN + трансформер + ASPP + канальное внимание) получает DSC до 0.93 на комбинированном наборе данных, что лучше, чем у Swin U-Net и TransUNet [15].

Обзор 67 статей по нейровизуализации, написанных с использованием трансформеров, выявил, что 85 % из них используют гибридные модели CNN-трансформер, которые, как правило, превосходят как автономные CNN, так и чистые трансформеры. в частности, однако,

за счёт более высоких вычислительных затрат и более высокого риска переобучения на ограниченных наборах данных [16].

### **III. СИСТЕМА ОБУЧЕНИЯ И ОСНОВНЫЕ ТЕХНИЧЕСКИЕ ВОПРОСЫ**

Сегментация опухолей мозга с использованием глубокого обучения зависит не только от сетевой архитектуры. на эффективность и точность алгоритмов значительно влияет организация обучения. Важными компонентами процесса являются детальная предварительная подготовка данных, методы борьбы с сильным дисбалансом классов, а также современные техники оптимизации.

#### *А. Pipeline данных*

- Предварительная обработка: Предварительная обработка МРТ включает в себя применение операции удаления черепа, перерисовку с интерполяцией до равномерного расстояния между voxels и нормализацию интенсивности либо коррекцию поля смещения. Это существенно увеличивает различия между центрами [17].
- Обучение по фрагментам или целым срезам: Обучение по фрагментам позволяет уменьшить требования к памяти, увеличить эффективный размер выборки и частично помочь в борьбе с дисбалансом классов. Перекрывающиеся или многомасштабные фрагменты фиксируют как локальные детали, так и более широкий окружающий контекст.
- Расширение данных (отражение, поворот, масштаб, изменение яркости, упругая деформация) весьма необходимо при малочисленности меченых данных [17]. Целевое или «избирательное» расширение и 3D-возмущения упругости/яркости сильно положительно влияют на чёткость Dice для 2D/3D U-Nets.

#### *В. Дисбаланс меток и функции потерь*

Образцов (вокселей) опухолевой ткани катастрофически меньше, чем нормальной, что и создаёт дисбаланс классов. Среди возможных средств решения проблемы можно отметить двухфазное обучение, cascade CNN, выборку фрагментов с прицелом на опухоль, а также применение критериев Dice и фокальных потерь, направленных на выделение классов меньшинства [17].

#### *С. Оптимизация и метаэвристики*

Стандартные оптимизаторы, такие как Adam и SGD с импульсом, по-прежнему наиболее популярны, но вдохновлённые природой метаэвристики, такие как PSO, GA, Grey Wolf Optimizer, становятся все более популярными в задачах настройки гиперпараметров, архитектур и внимания. Эти алгоритмы зачастую приводят к повышению индексов Dice и Jaccard в мультимодальных задачах [18].

Для построения надежной системы сегментации опухолей головного мозга не хватает одной архитектуры модели. Важны стандартизированная предварительная обработка, не чрезмерно изощрённая, обучение на патчах или срезах с серьёзной аугментацией, чёткие

планы по борьбе с дисбалансом классов как в выборке, так и в функциях потерь, а также грамотно подобранные оптимизаторы (с растущим интересом к настройке метаэвристик) — всё это неотъемлемые элементы эффективных тренировочных конвейеров.

#### IV. КЛИНИЧЕСКОЕ ПРИМЕНЕНИЕ И ПРОБЛЕМЫ

##### A. Клиническое значение

Сегментация, основанная на глубоких обучении, содействует в планировании лечения, оценке объёмов, поиске мишеней для лучевой терапии, мониторинге ответа на лечение или псевдопрогрессирования при глиомах [17]. Исследователи изучают возможности провести соответствие фенотипов визуализации и молекулярных подтипов и прогнозов в рамках радиомики и геномики [17].

##### B. Основные ограничения

- **Обобщаемость:** эффективность зачастую падает на незнакомых сканерах/в незнакомых больницах из-за смены домена; адаптация и согласование доменов являются актуальными приоритетами [17][18].
- **Недостаток данных:** аннотированные 3D-объёмы МРТ очень дороги; исследователи смотрят на переносное обучение, полунаблюдаемое обучение, синтетические данные [17][18].
- **Вычислительные затраты:** 3D-CNN и трансформеры могут быть жадными к ресурсам; исследователи смотрят на облегчённые архитектуры, методы обрезки/квантования, чтобы обеспечить возможность работы в реальном времени или на ограниченных ресурсах [18].
- **Интерпретируемость и доверие:** DL часто является «чёрным ящиком». Методы XAI, такие как Grad-CAM, SHAP, LIME, используются для демонстрации того, какие области данных влияют на прогнозы, и повышения доверия врачей к системе [18].

В табл. I ниже методология, производительность, основные плюсы и минусы нескольких недавних статей суммированы.

ТАБЛИЦА I. Сводка некоторых статей по сегментации опухолей головного мозга: методология, производительность и критические замечания

Источник	Основная методология	Типичная производительность	Сильные стороны	Слабые стороны / Проблемы
[19]	Важнейшая методология, на основе которой осуществлялись, заключалась в использовании 2D-CNN со смешанным локальным глобальным путями, полностью свёрточном последнем слое для ускорения, двухступенчато	BraTS-2012: Dice 0.81 (в целом), 0.72 (основной блок), 0.58 (усиливающий блок); на тот момент она была передовой технологией и работала примерно в 30–40 раз быстрее, чем CNN с	первая полностью автоматическая CNN для BraTS; применение локального и глобального контекста; эффективный свёрточный «FC»-слой; явная стратегия решения	только 2D (ограниченна я поддержка 3D); основана на классических блоках CNN; по производительности не дотягивает до более поздних моделей U-Net/3D, особенно в

Источник	Основная методология	Типичная производительность	Сильные стороны	Слабые стороны / Проблемы
	м обучении для разрешения проблемы дисбаланса меток; каскадной структуре, использующей выход первой Cnn в качестве входа для второй.	блоками-патчами.	проблемы дисбаланса классов.	плане улучшения распознавания опухоли.
[20]	Каскадная CNN (C-ConvNet) на обрезанных областях интереса (ROI) для уменьшения вычислений, два параллельных пути, фиксирующих локальные и глобальные признаки, новейший механизм «Distance-Wise Attention» (DWA) для подбора веса признаков в зависимости от расстояния до центра опухоли/головного мозга.	Brats2018: Обычное значение коэффициента Dice: 0.9203 (whole), 0.9113 (enhancing), 0.8726 (core).	Хорошее значение Dice при относительно простой архитектуре. Обрезка, каскадирование минимизируют переобучение и экономят время выполнения. Dwa улучшает точность по границам.	По-прежнему на конкретном наборе данных (только Brats); по-прежнему акцент на пространстве вблизи расстояния, а не на текстуре. По-прежнему не изучены возможности обобщения на другие патологии сканеры.
[7]	Модель сегментации, основанная на U-Net с кодировщиком EfficientNetB4 для иерархического извлечения признаков и декодировщиком, состоящим из многоуровневых резидуальных блоков с механизмом внимания (1×1, 3×3, 5×5 convolutions) и остаточными связями, для уточнения масок опухолей с высоким разрешением на данных МРТ опухолей мозга из Figshare	Среди всех вариантов EfficientNet модель B4 продемонстрировала лучшую сегментацию с показателями Dice 0.9339, IoU 0.8795, а предлагаемый метод опередил все базовые модели на Figshare, получив точность 99.79%, Dice 0.9339, IoU 0.8795, высокую точность (0.9657) и полноту (0.9103)	Хорошее значение Dice	Оценка проводилась исключительно на одном наборе данных Figshare, поэтому возможность обобщения результатов на другие центры и методы визуализации не подтверждена. Модель предназначена исключительно для бинарной сегментации (опухоль против фона). Перекрестная валидация не проводилась из-за высоких вычислительных затрат. Отмечены

Источник	Основная методология	Типичная производительность	Сильные стороны	Слабые стороны / Проблемы
				ограничения при работе с опухолями с низкой контрастностью и сложной формой, а результаты сравнения с другими моделями по скорости вывода не представлены систематически.

## V. ЗАКЛЮЧЕНИЕ

Данное глубокое обучение, основанное на 2D/3D-архитектурах U-Net, осуществляемое путём применения CNN, по-прежнему остаётся основополагающим подходом к сегментации опухолей головного мозга на мультимодальных МРТ-изображениях, стабильно демонстрируя высокие показатели коэффициента Дайса на эталонных наборах данных. Варианты U-Net с улучшенными модулями кодирования, внимания или Inception могут слегка повысить прагматическую эффективность, однако в данной области наблюдается всё более настойчивый и заметный сдвиг в сторону гибридных моделей, сочетающих трансформеры или графовые нейронные сети. Эти архитектуры повышают точность за счёт превосходного моделирования глобального контекста, но влекут за собой значительные компромиссы в плане сложности, вычислительных затрат и требований к данным. Несмотря на то, что эти модели становятся всё более точными с клинической точки зрения, их нельзя использовать в клинической практике до тех пор, пока не будут решены проблемы с обобщением на межцентровую вариабельность, эффективностью данных, объяснимостью моделей и созданием оптимизированных сквозных клинических конвейеров.

## СПИСОК ЛИТЕРАТУРЫ

- [1] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, "Cancer statistics, 2023," *CA a Cancer Journal for Clinicians*, vol. 73, no. 1, pp. 17–48, Jan. 2023, doi: 10.3322/caac.21763.
- [2] Z. U. Abidin, R. A. Naqvi, A. Haider, H. S. Kim, D. Jeong, and S. W. Lee, "Recent deep learning-based brain tumor segmentation models using multi-modality magnetic resonance imaging: a prospective survey," *Frontiers in Bioengineering and Biotechnology*, vol. 12, p. 1392807, Jul. 2024, doi: 10.3389/fbioe.2024.1392807.
- [3] R. Ranjbarzadeh, A. B. Kasgari, S. J. Ghousechi, S. Anari, M. Naseri, and M. Bendeche, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Scientific Reports*, vol. 11, no. 1, p. 10930, May 2021, doi: 10.1038/s41598-021-90428-8.
- [4] R. Yousef et al., "U-Net-Based Models towards Optimal MR Brain Image Segmentation," *Diagnostics*, vol. 13, no. 9, p. 1624, May 2023, doi: 10.3390/diagnostics13091624.
- [5] F. F. Alkhalid and N. Z. Salih, "Implementation of biomedical segmentation for brain tumor utilizing an adapted U-net model," *Computers in Biology and Medicine*, vol. 194, p. 110531, Jun. 2025, doi: 10.1016/j.compbiomed.2025.110531.
- [6] S. Ghosh, A. Chaki, and K. Santosh, "Improved U-Net architecture with VGG-16 for brain tumor segmentation," *Physical and Engineering Sciences in Medicine*, vol. 44, no. 3, pp. 703–712, May 2021, doi: 10.1007/s13246-021-01019-w.
- [7] Preetha R., Jasmine Pemeena Priyadarsini M., Nisha J. S., "Brain tumor segmentation using multi-scale attention U-Net with EfficientNetB4 encoder for enhanced MRI analysis," *Scientific Reports*, vol. 15, no. 1, p. 9914, Mar. 2025, doi: 10.1038/s41598-025-94267-9.
- [8] E. K. Rutoh, Q. ZhiGuang, J. C. Bore-Norton, and N. Bahadar, "ABI-NET: Attention-Based Inception U-Net for brain tumor segmentation from multimodal MRI images," *IEEE Access*, vol. 13, pp. 134898–134916, Jan. 2025, doi: 10.1109/access.2025.3585926.
- [9] J. Liu, U. A. Bhatti, J. Zhang, Y. Zhang, and M. Huang, "EF-VPT-NET: Enhanced Feature-Based Vision Patch Transformer Network for accurate brain tumor segmentation in magnetic resonance imaging," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, pp. 1–14, Jan. 2025, doi: 10.1109/jbhi.2025.3526976.
- [10] W. Wang, C. Chen, M. Ding, H. Yu, S. Zha, and J. Li, "TransBTS: Multimodal brain tumor segmentation using transformer," in *Lecture notes in computer science*, 2021, pp. 109–119. doi: 10.1007/978-3-030-87193-2\_11.
- [11] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu, "SWIN UNETR: SWIN Transformers for Semantic Segmentation of Brain Tumors in MRI images," in *Lecture notes in computer science*, 2022, pp. 272–284. doi: 10.1007/978-3-031-08999-2\_22.
- [12] I. Aboussaleh, J. Riffi, K. E. Fazazy, A. M. Mahraz, and H. Tairi, "3DUV-NetR+: A 3D hybrid semantic architecture using transformers for brain tumor segmentation with MultiModal MR images," *Results in Engineering*, vol. 21, p. 101892, Feb. 2024, doi: 10.1016/j.rineng.2024.101892.
- [13] K. Jagani, M. A. Hanif, J. A. A. Jothi, and M. Shafique, "CA-U-Net: Context Aware U-Net For Brain Tumor Segmentation," *IEEE Xplore*, pp. 1–7, Jun. 2025, doi: 10.1109/ijcnn64981.2025.11228675.
- [14] M. Zhang et al., "Augmented Transformer network for MRI brain tumor segmentation," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 1, p. 101917, Jan. 2024, doi: 10.1016/j.jksuci.2024.101917.
- [15] M. N. Nagib, R. Pervez, A. A. Nova, H. R. Nabil, Z. Aung, and M. F. Mridha, "TUSEGNET: A Transformer-Based and Attention-Enhanced architecture for brain tumor segmentation," *IEEE Open Journal of the Computer Society*, vol. 6, pp. 750–761, Jan. 2025, doi: 10.1109/ojcs.2025.3569758.
- [16] M. Iratni et al., "Transformers for Neuroimage Segmentation: Scoping Review," *Journal of Medical Internet Research*, vol. 27, p. e57723, Jan. 2025, doi: 10.2196/57723.
- [17] A. Verma and A. K. Yadav, "Brain tumor segmentation with deep learning: Current approaches and future perspectives," *Journal of Neuroscience Methods*, vol. 418, p. 110424, Mar. 2025, doi: 10.1016/j.jneumeth.2025.110424.
- [18] A. Zaman et al., "Challenges, optimization strategies, and future horizons of advanced deep learning approaches for brain lesion segmentation," *Methods*, vol. 239, pp. 140–168, Apr. 2025.
- [19] M. Havaei et al., "Brain tumor segmentation with Deep Neural Networks," *Medical Image Analysis*, vol. 35, pp. 18–31, May 2016, doi: 10.1016/j.media.2016.05.004.
- [20] R. Ranjbarzadeh, A. B. Kasgari, S. J. Ghousechi, S. Anari, M. Naseri, and M. Bendeche, "Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images," *Scientific Reports*, vol. 11, no. 1, p. 10930, May 2021.