

3D-СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ ОПУХОЛЕЙ ГОЛОВНОГО МОЗГА НА СНИМКАХ МРТ

Закиев Я.Т.¹, Спицын В.Г.^{2,3}

¹НИ ТГУ, ИПМКН, аспирант, ya@stud.tsu.ru

²НИ ТГУ, ИПМКН, д.т.н., профессор каф. ТОИ

³НИ ТПУ, ИШИР, д.т.н., профессор ОИТ

Аннотация

Была решена задача семантической сегментации снимков МРТ опухолей головного мозга в 3D-измерении при помощи глубокой сверточной нейронной сети архитектуры DynU-Net. Для разработки использовался фреймворк Monai и набор данных Brats 2023, к которому был применен ряд операций аугментации.

Ключевые слова: изображения МРТ, классификация, сегментация, сверточные нейронные сети, Monai.

Введение

По данным межконтинентального ракового регистра, объединяющего данные из 86 раковых регистров 5 континентов, в 2022 году в мире было зарегистрировано 322 000 новых случаев опухолей головного мозга и ЦНС [1]. Данная мировая статистика собирается Международным агентством по изучению рака при участии Всемирной Организации Здравоохранения. В России первичные опухоли головного мозга ежегодно обнаруживают примерно у 30 тыс. человек. Наибольшее число заболевших приходится на возраст от 40 до 54 лет (30,8%) и от 55 до 69 лет (31,4%) [2]. С целью диагностики в настоящее время производится ручная сегментация МРТ изображений.

Согласно статье [3] основные проблемы ручной сегментации МРТ изображений опухолей головного мозга на сегодняшний день в том, что она:

1. Требуется больших затрат времени радиологов высокой квалификации.

2. Даже внутри одного единственного исследовательского центра нет единого стандарта границ сегментации опухоли, что говорит о субъективизации, также есть определенный процент ошибок. В целом этот пункт можно охарактеризовать, как человеческий фактор, который играет определенную роль в диагностике опухолей головного мозга. Также отмечается, что постепенно методы машинного и глубинного обучения могут стать стандартом для этих целей [3].

Сегментация опухолей в 3D – измерении позволяет проводить количественную оценку анатомических или патологических поражений, что дает более точную информацию по сравнению с качественными 2D измерениями [4].

В соответствии с этим, целью данной работы является применение глубокого обучения, способного производить 3D сегментацию опухолей головного мозга, с получением модели на современном уровне показателей.

Разработанный подход Monai DynU-Net

В работе использовался фреймворк Monai [5]. MONAI (Medical Open Network for AI) – это фреймворк с открытым исходным кодом, разработанный специально для удобной работы с медицинскими изображениями. Он построен на основе PyTorch и предлагает набор инструментов для глубокого обучения в области медицинской визуализации. Основные возможности MONAI это работа с медицинскими данными, он поддерживает стандартные форматы NIfTI, DICOM и др. Monai имеет встроенную реализацию предобработки изображений: функции для нормализации, аугментации, ресамплинга и работы с масками. Фреймворк содержит многие архитектуры сетей, включая популярные модели, такие как UNet, DynU-Net, SwinUNETR и др. Monai совместим с PyTorch Lightning, он упрощает процессы

обучения, валидации и тестирования. Monai включает инструменты оценки метрики (DSC, IoU), функции потерь (Dice Loss, Focal Loss и др.). Он также работает с GPU, TPU, многопроцессорными системами и облачными средами.

Для реализации модели была выбрана архитектура DynU-Net [5]. DynU-Net – это адаптация архитектуры U-Net, разработанная специально для библиотеки MONAI. Она предназначена для задач сегментации изображений, особенно в медицинской области. Ключевыми особенностями архитектуры DynU-Net являются динамические блоки, это означает, что структура модели может адаптироваться (или "настраиваться") в зависимости от входных данных или параметров задачи. Это делает архитектуру более универсальной и эффективной для различных условий, задач или размеров данных. Параметры нейронной сети задаются через удобный API. Задается размерность, число входных и выходных каналов.

Понижение частоты дискретизации происходит с помощью Strided Convolutions: в отличие от стандартной U-Net, которая использует max-pooling для уменьшения разрешения, DynU-Net применяет strided convolutions (свёрточные слои с шагом больше 1). Это позволяет сети обучаться более эффективному понижению частоты дискретизации, точнее регулировать степень уменьшения разрешения и лучше сохранять пространственные особенности данных. Адаптированный Residual Block: DynU-Net модифицирует residual block (остаточный блок) с помощью дополнительного свёрточного слоя. Этот слой гарантирует, что размер входных данных с понижением частоты дискретизации слоя соответствует размеру выходных данных residual block'a, обеспечивая корректное сложение residual connection (остаточного соединения).

Архитектура DynU-Net допускает изменение количества слоёв и размеров фильтров (ядер свёртки). Стандартная конфигурация использует 6 слоёв с размером фильтров [32, 64, 128, 256, 320, 320]. Размеры шагов свёртки для понижения частоты дискретизации – [1, 2, 2, 2, 2, [2, 2, 1]]. Upsampling выполняется в обратном порядке. Также можно задавать размер ядра kernel size и шаг stride. DynU-Net сохраняет skip connections (пропускные соединения) из U-Net, позволяя передавать информацию с ранних слоёв понижения частоты дискретизации на соответствующие слои повышения частоты дискретизации. Это помогает восстановить мелкую детализацию при увеличении разрешения изображения. DynU-Net поддерживает различные типы нормализации:

- Batch Normalization. Устраняет смещения между мини-батчами. Полезно для обучения на больших наборах данных.
- Instance Normalization. Нормализует каждый пример в отдельности. Часто применяется в медицинской обработке изображений.
- Group Normalization. Делит каналы на группы и нормализует их, хорошо подходит для небольших батчей.

В DynU-Net есть поддержка глубокой супервизии deep_supervision, позволяющей модели генерировать прогнозы на промежуточных уровнях (глубоких слоях). В целом, DynU-Net – это гибкая и мощная архитектура для сегментации, построенная на основе U-Net. Использование strided convolutions и адаптированного residual block'a, наряду с возможностью конфигурирования глубины и размеров фильтров, делает её особенно эффективной для анализа медицинских изображений.

В дополнение к вышесказанному, стоит отметить следующие аспекты: DynUNet доступна в библиотеке MONAI, предоставляющей инструменты для глубокого обучения в медицинской визуализации. DynUNet часто сравнивают с nnU-Net, ещё одной популярной архитектурой для сегментации медицинских изображений. DynU-Net можно рассматривать как реализацию архитектуры, подобной nnU-Net, в рамках MONAI [6]. DynU-Net успешно применяется для сегментации различных органов на КТ и МРТ изображениях. На данный момент DynU-Net активно развивается в MONAI и является интересной архитектурой для исследований задачи сегментации медицинских изображений.

Работа с датасетом

Разрабатываемая модель обучалась на датасете **BRATS2023 (Brain Tumor Segmentation Challenge)** часть 1. Он включает МРТ-изображения головного мозга с разметкой опухолей, предназначенные для задач сегментации. Состоит из различных типов сканов, таких как **T1**, **T1Gd**, **T2** и **FLAIR**, и включает аннотированные маски для четырех категорий: некротической опухоли **NCR** (necrotic tumor core), усиливающейся опухоли (enhanced tumor) **ET**, перитуморального отека (peritumoral edematous, edema) **ED** и фона (background) **Back** – здоровых тканей. Все данные представлены в формате NIfTI, с размерами изображений 240x240x155 вокселей. Была применена нормализация и аугментация:

- нормализована интенсивность пикселей изображения, за счёт приведения их в диапазон [0, 1]. Включена опция для нормализации по каналам и для игнорирования нулевых значений.
- случайное обрезание изображения и метки в заданный размер области интереса (ROI) [96, 96, 96], без случайного изменения размера изображения.
- случайные перевороты изображения и меток по осям.
- случайное изменение интенсивности изображения.
- случайное смещение интенсивности изображения на небольшое значение (до 0.1).

Сравнение результатов

Результаты показывают, что DynU-net не смотря на небольшое количество параметров показывает себя конкурентной в сравнении с другими современными архитектурами. Из рисунка 1 следует, что функция потерь Average Dice и метрика Mean Dice для всех классов, соответственно и средний Mean Dice по всем трем классам, стабильно сходятся. Расстояние Hausdorff имело затухающие биения для всех классов, но также сходилось.

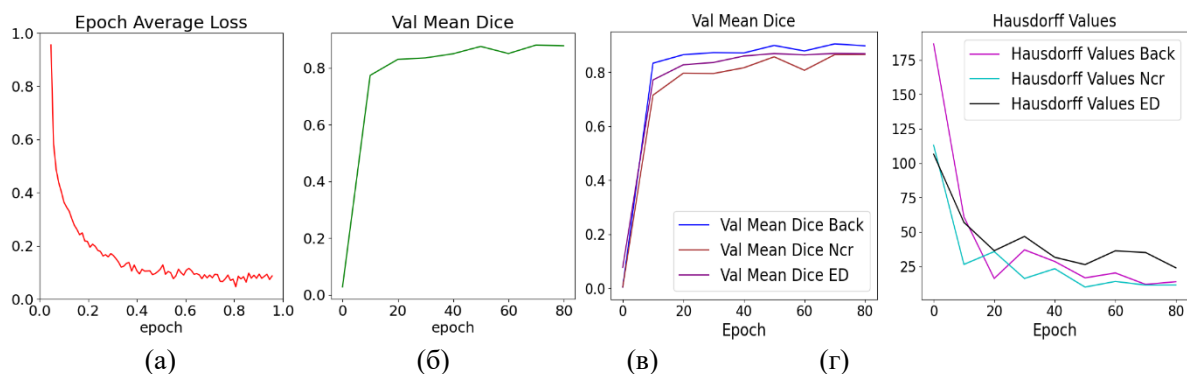


Рис. 1. Графики обучения DynU-Net: а – Average Dice loss, б – Mean Dice средний по трем классам, в – Mean Dice, г – Hausdorff Values – по трем классам

На рисунке 2 отображена матрица ошибок предсказаний Dyn U-net, видно, что модель иногда путает третий класс ET с остальными, ложных пропусков FN по классу ED больше: $52 + 956 = 1008$ против $307 + 295 = 602$, тогда как ложных срабатываний FP $17 + 295 = 312$ против $792 + 956 = 1748$. Можно сделать вывод модель иногда плохо обрисовывает зону отека и иногда плохо различает отек и зону усиливающейся опухоли.

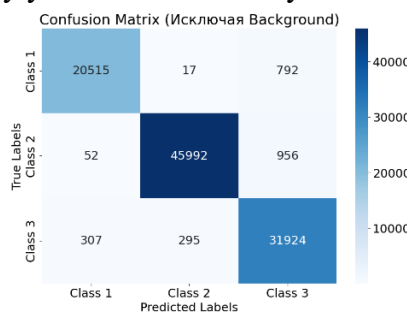


Рис. 2. Confusion matrix для трех классов: NCR, ED, ET

Класс номер 2 ED (Edema) имеет самую низкую метрику Mean Dice, что видно на результатах сегментации на рисунке 3.

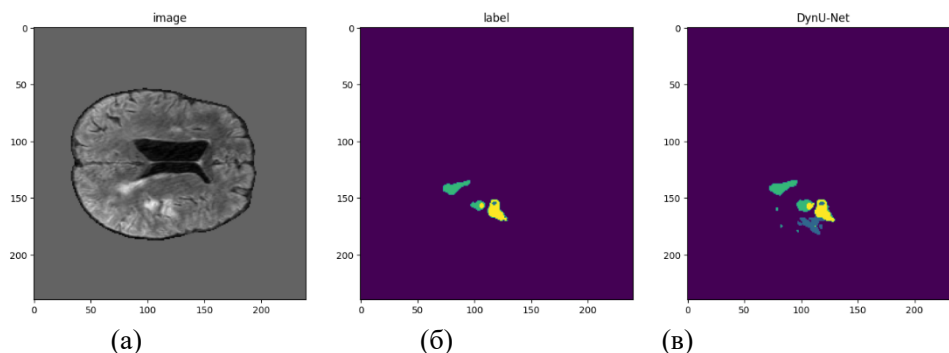


Рис. 3. Результаты применения сегментации: а – оригинал, б – оригинальная, в – предсказанная маска

С другой стороны ее значение на хорошем уровне относительно других работ **Brats 2023** (таблица 2, отёк может быть диффузным, не иметь чётких границ. Тогда как расстояние Hauardsoff для класса номер 2 **ED** (перитуморального отёка) показывает худшие результаты. Размытые края отёка могут затруднять точную локализацию. Mean Dice для класса 3 **ЕТ** (усиливающейся опухоли) и остальные метрики в сравнении с работами **Brats 2023** на отличном уровне. Метрика Hauardsoff по остальным классам находятся на достойном, либо лучшем уровне. Общее количество параметров DynU-Net: 31411844, все параметры обучаемые.

Таблица 2. Сравнение результатов семантической сегментации DynU-Net с лидербордом Brats 2023 [10]

Авторы	Lesion Wise Dice ET	Lesion Wise Dice TC	Lesion Wise Dice WT	Lesion Wise Hausdorff ET	Lesion Wise Hausdorff TC	Lesion Wise Hausdorff95 WT
@s7zuqer	0.866	0.89	0.909	10.594	7.228	11.086
@simis	0.857	0.876	0.908	13.649	12.825	11.27
@qian8	0.856	0.868	0.904	14.07	13.778	10.735
@qian6	0.856	0.868	0.905	14.07	13.778	12.563
@qian5	0.856	0.868	0.907	14.07	13.783	11.797
@qian9	0.856	0.869	0.904	14.068	13.775	10.736
@qian7	0.856	0.869	0.904	14.068	13.775	12.567
@qian6	0.856	0.869	0.907	14.068	13.78	11.08
@simis	0.856	0.869	0.904	14.069	12.08	12.565
@kai_kai	0.856	0.869	0.905	14.069	12.08	12.693
Наши	0.952	0.972	0.946	14.071	9.849	45.126

Результаты предсказаний DynU-Net отдельно по тканям классов **NC**, **ED**, **ЕТ** представлены в таблице 2. Высокое значение расстояние Hauardsoff у перитуморального отека **ED** объясняет высокое значение Hauardsoff с перцентилем 95, принятам для Brats, для объема всей опухоли WT.

Таблица 2. Метрики семантической сегментации DynU-Net для классов NC, ET, ED

Метрика	NC	ED	ЕТ
Lesion Wise Dice	0.944	0.892	0.952
Lesion Wise Hausdorff	9.434	56.833	14.071

Далее сравним результаты с хорошо зарекомендовавшими себя трансформерными архитектурами [9] таблица 3.

Таблица 3. Сравнение с результатами других работ, посвященных сегментации опухолей мозга [9]

Архитектура	Параметры	Lesion Wise Dice ET	Lesion Wise Dice TC	Lesion Wise Dice WT
nnFormer	150.5	91.3	81.8	86.0
SegFormer[9]	4.5	89.9	74.2	82.2
EHETP	92.49	78.9	58.5	76.1
TransBT	-	77.9	57.4	73.5
CoTr	41.9	74.6	55.7	74.8
CoTr без кодировщика CNN	-	71.2	52.3	69.8
TransUNet	96.07	70.6	54.2	68.4
SETR MLA	310.5	69.8	55.4	66.5
SETR PUP	318.31	69.6	54.9	67.0
SETR NUP	305.67	69.7	54.4	66.9
DynU-Net	31,41	89.2	0.909	94.6

В данной работе авторы указывают, что архитектуры на сверточных сетях, обладая индуктивным смещением, традиционно показывают высокую конкурентность при обучении с нуля и получают гибридную архитектуру. В этом состоит их отличие от трансформерных архитектур, которым требуются большие наборы данных, недоступные в задачах медицинской обработки изображений. Особенно важно, что трансформерные архитектуры затратны по количеству параметров, что очень заметно в задачах 3D сегментации, так как объемы массивов увеличиваются и в настоящее время исследования направлены на гибридные архитектуры, которые бы обладали меньшим числом параметров и плюсами CNN.

Авторы описывают модель сегментации, основанную на иерархическом ViT, которая эффективно обрабатывает медицинские изображения. Она использует модуль внимания, сохраняет локальную непрерывность за счет модуля эмбединга перекрывающихся патчей и имеет легкую архитектуру из-за отказа от тяжелого декодера в пользу декодера all-MLP. Модель имеет низкую сложность (4.5 млн параметров, 17 GFLOPS), но сохраняет высокую производительность, что делает её конкурентоспособной с лучшими моделями (SOTA) и она не нуждается в предварительном обучении.

Авторы [11] описывают DCDynU-Net – это усовершенствованная модель сегментации медицинских изображений (например, МРТ мозга), основанная на DynU-Net. Она использует глубокий супервизор и блоки 3D CBam для улучшения кодирования признаков. Декодер включает многоголовый блок внимания, объединённый с механизмом "глубокого супервизора", эффективно комбинирующий информацию с предыдущих этапов декодирования и последних слоёв, для повышения качества сегментации.

Суммарно по метрикам, и по числу параметров 31,4 млн. наша архитектура показывает себя на хорошем уровне. Данная архитектура может стать отличной основой для исследования в сторону гибридных архитектур. Авторы [12] используют разработку DeepDynU-Net, сочетающую в себе преимущества глубокого супервизора и механизмов внимания, применяя их к архитектуре DynU-Net. Обученный на Brats 2018 совместно с DynU-Net обученным на iSeg19 он применяется в оригинальном подходе, позволяющим улучшить качество сегментации, для объединения контекстной информации Context Feature Fusion о составе мозга с признаками опухоли.

Заключение

В ходе проделанной работы была получена модель архитектуры DynU-Net с помощью современного фреймворка для обработки медицинских изображений Monai. Результаты семантической сегментации показывают хороший уровень относительно других работ на

соревновании Brats 2023 и некоторых современных трансформерных решений. По количеству параметров и метрикам разработанная модель оказывается выгодной в сравнении с последними, что характеризует ее как хорошую базу для дальнейших исследований в направлении развития гибридных архитектур.

Список использованных источников

1. Adalberto Filho M. et al. Cancers of the brain and central nervous system: global patterns and trends in incidence // *Journal of neuro-oncology*. – 2025. – Т. 172. – №. 3. – С. 567-578.
2. Опухоли головного мозга [Электронный ресурс]. – URL: pirogov-center.ru/specialist/diseases/detail.php?ID=479 (дата обращения: 10.04.2025)
3. Далечина А.В. Современные подходы к диагностике и лечению глиобластомы / А.В. Далечина, М.Г. Беляев, А.Н. Тюрина и др. // АО «Деловой центр нейрохирургии» (Центр «Гамма-нож»), Москва, Россия; Сколковский институт науки и технологий, Москва, Россия; НМИЦ нейрохирургии им. акад. Н. Н. Бурденко, Москва, Россия. © Коллектив авторов, 2019.
4. Dupont C., Betrouni N., Reyns N. et al. On image segmentation methods applied to glioblastoma: state of art and new trends // *IRBM, Elsevier Masson*. – 2016. – Vol. 37, no. 3. – P. 131-143. DOI: 10.1016/j.irbm.2016.04.002.
5. Ruffle J. K., Mohinta S., Gray R., Hyare H., Nachev P. Brain tumour segmentation with incomplete imaging data // *Brain Communications*. – 2023. – Vol. 5, no. 2. DOI: 10.1093/braincomms/fcad118.
6. Cardoso M. J. et al. MONAI: An open-source framework for deep learning in healthcare // *arXiv preprint arXiv:2211.02701*. – 2022. URL: arxiv.org/abs/2211.02701.
7. MONAI DynUNet documentation // MONAI. – URL: docs.monai.io/en/1.3.0/_modules/monai/networks/nets/dynunet.html (дата обращения: 06.11.2024).
8. DynUNet pipeline tutorial // Project-MONAI. – URL: https://github.com/Project-MONAI/tutorials/tree/main/modules/dynunet_pipeline (дата обращения: 10.12.2024).
9. Perera S., Navard P., Yilmaz A. Segformer3d: an efficient transformer for 3D medical image segmentation // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. – 2024. – С. 4981–4988.
10. BraTS 2023 Challenge. – URL: synapse.org/Synapse:syn51156910/tables/ (дата обращения: 05.02.2024).
11. Son V. T. H. et al. DCDynUnet: Deep Supervision Attention Context for Brain Segmentation // *International Conference on Advances in Information and Communication Technology*. – Cham: Springer Nature Switzerland, 2023. – С. 158–165.
12. Thanh N.T.N., Huy N.Q., Hyung K.S. Brain tumor segmentation based on deep supervision and context feature fusion.