



МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М. В. ЛОМОНОСОВА
ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ
КАФЕДРА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Айрапетьянц Каринэ Арсеновна
**Обзор методов дифференциальной диагностики глиальных опухолей
по данным динамических ПЭТ - исследований**

КУРСОВАЯ РАБОТА

Научный руководитель:
Малоян Нарек Гагикович

Москва, 2022

Оглавление

Введение	3
1 Методы	7
1.1 Методы сегментации и классификации	7
Список литературы	17

Введение

Термин «первичные опухоли ЦНС» (ПО ЦНС) объединяет различные по гистологическому строению, злокачественности и клиническому течению опухоли, общим для которых является происхождение из тканей, составляющих центральную нервную систему и ее оболочки. [1]

Опухоли головного мозга (ОГМ)- это группа различных внутричерепных новообразований (доброкачественных и злокачественных), возникающих вследствие запуска процесса аномального неконтролируемого деления клеток, которые в прошлом являлись нормальными составляющими различных тканей головного мозга, или возникающих вследствие метастазирования первичной опухоли, находящейся в другом органе. На возникновение новообразований влияют различные факторы: некоторые внутренние редкие дефекты генов, внешние воздействия - ультрафиолетовые лучи или рентгеновское излучение, химические вещества, инфекции, - приводят к спонтанному появлению мутации, что повышает риск развития онкологических заболеваний. Опухоли головного мозга с разными частотами затрагивают все возрастные категории населения. Прогноз заболевания в большинстве случаев остается неблагоприятным. Продолжительность жизни больных с ОГМ значительно варьируется в зависимости от типа новообразования, составляя в среднем от 1 года до 3-7 лет. Подобные новообразования встречаются с частотой от 5 до 7,5 случаев на 100 тысяч населения. Поэтому одним из наиболее значимых приоритетных направлений современной медицины является совершенствование существующих методов диагностики и терапии ОГМ. [2]

Одной из наиболее часто встречающихся злокачественных патологий центральной нервной системы является глиальная опухоль.

Глиомы – это собирательный термин, который объединяет все диффузные астроцитарные и олигодендроглиальные опухоли, а также другие виды – пилоидную астроцитому, субэпендимарную гигантоклеточную астроцитому, астробластому и другие опухоли, исходящие из клеток глии. [1]

Опухоли ЦНС очень разнообразны. Их классифицируют по локализации, гистологическому типу, степени злокачественности.

Классификация глиом [1]:

- 1) Астроцитомы – опухоль, развивающаяся из астроцитарной части глии и представленная астроцитами. Может локализоваться как в больших

полушариях мозга, так и в мозжечке, а также в стволе головного мозга и спинном мозге. Различают астроцитомы низкой и высокой степени злокачественности.

- 2) Олигодендроглиома и олигоастроцитомы – опухоли, преимущественно состоящие из олигодендроцитов.
- 3) Глиоматоз головного мозга – это диффузное поражение глиомой структур головного мозга (более 3-х анатомических областей больших полушарий, обычно с переходом через мозолистое тело и с перивентрикулярным распространением).
- 4) Глиомы ствола головного мозга. На разных уровнях поражения ствола головного мозга встречаются различные глиальные опухоли. Часть этих опухолей носит доброкачественный характер и может не прогрессировать без лечения в течение всей жизни человека. Другие характеризуются, напротив, агрессивным течением с ограниченными возможностями специализированной помощи этим больным.
- 5) Глиомы спинного мозга. Как правило, диффузные интрамедуллярные опухоли, поражающие различные уровни спинного мозга.
- 6) и т.д.

Методы обследования глиальных опухолей:

- 1) Компьютерная томография с контрастным усилением (КТ) - это специальный метод, в котором используется рентгеновское излучение. С его помощью тело человека послойно просвечивают рентгеновскими лучами.
- 2) Магнитно-резонансная томография с контрастным усилением (МРТ) - метод визуализации, основанный на резонансе атомов водорода в организме человека на магнитное поле, создаваемое томографом.
- 3) Позитронно-эмиссионная томография (ПЭТ) - технология визуализации, основанная на количественной и качественной оценке биохимических процессов, происходящих в тканях *in vivo*.

- 4) Сцинтиграфия - используется для оценки функционирования различных органов и тканей. Такие методы диагностики, как рентген, УЗИ, КТ или МРТ ориентированы на выявление структурных изменений в тканях организма, и не всегда способны различить болезнь на ранних её стадиях, когда отклонения проявились на уровне биохимических изменений в тканях. В это время приходит на помощь сцинтиграфия, которую поэтому и называют молекулярной диагностикой
- 5) Неврологическое исследование, которое обязательно включает в себя офтальмологическую проверку остроты зрения, глазного дна и полей зрения
- 6) Ангиография - класс методов контрастного исследования кровеносных сосудов.

На данный момент «золотым стандартом» в диагностике объемных образований головного мозга, определении степени злокачественности, тактики лечения и прогноза заболевания является магнитно-резонансная томография (МРТ) с контрастным усилением. Однако методы диагностики, направленные, главным образом, на оценку структурных изменений мозга, к которым относится и МРТ, обладают невысокой специфичностью в выявлении микроструктурных и метаболических перестроек в опухолевой ткани, что ограничивает раннюю диагностику глиального образования. Накопление МР-контрастного препарата не всегда напрямую коррелирует со степенью злокачественности опухоли, поэтому всё чаще и чаще для диагностики применяется ПЭТ-исследование совместно с КТ и МРТ исследованиями для уточнения результатов.

Ранняя диагностика опухоли способствует высоким результатам лечения. Заболевание нуждается в дифференцировке от гематомы внутри мозга, абсцесса, эпилепсии, прочих опухолевых процессов в центральной нервной системе, последствий инсульта.

В настоящее время все больше внимания уделяется автоматическим методам диагностики злокачественных новообразований (классификация, сегментация), разработанных на основе нейронных сетей, что в перспективе позволит ускорить процесс диагностики патологического процесса, сокращая время между проведением исследования и началом лечения, а также умень-

шить нагрузку на врачей и медработников, которые смогут более эффективно проводить лечение заболевания.

В данной работе проводится обзор различных нейросетевых методов, которые решают задачи сегментации, классификации и реконструкции медицинских изображений различных модальностей (ПЭТ, МРТ, КТ), которые могут применяться для автоматической диагностики злокачественных заболеваний. Сложность работы с медицинскими данными заключается в том, что зачастую качественных наборов данных очень мало, а для разметки нужно несколько специалистов - в некоторых случаях возникают спорные моменты и требуется дополнительное мнение. Рассмотренные модели показали хорошие результаты производительности в своей области применения, с помощью них можно эффективно обрабатывать, реконструировать медицинские данные, что отражает прогресс в разработке автоматических систем медицинской диагностики.

1. Методы

1.1. Методы сегментации и классификации

NormResSE-UNet3+

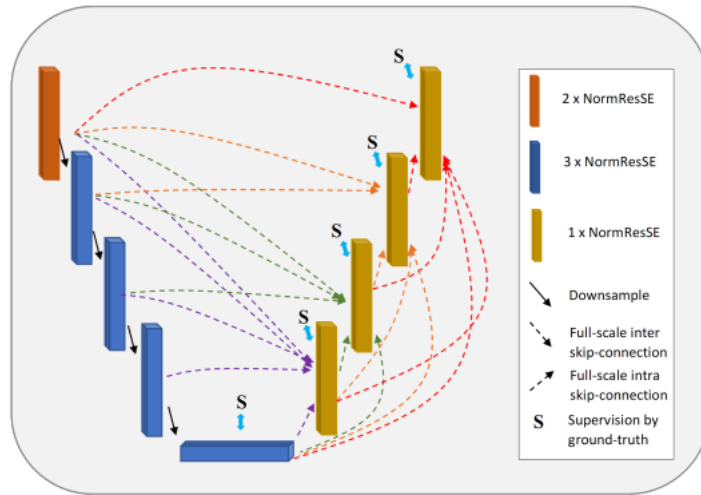
В работе [3] решается задача сегментации опухолей головы и шеи с помощью сверточной нейронной сети, а также задача предсказания выживаемости пациентов с помощью регрессионной модели.

Авторы предлагают производить сегментацию опухолей головы и шеи по ПЭТ/КТ изображениям, используя полномасштабную сеть архитектуры 3D UNet3++ [4] с механизмом, имитирующим когнитивное внимание. Предложенная модель, NormResSE-UNet3+ была обучена с гибридной функцией потерь, состоящей из Log Cosh Dice и Focal loss. Далее, предсказанные карты сегментации дополнительно уточняются с помощью механизма постпроцессинга - Conditional Random Fields, чтобы уменьшить количество ложноположительных ответов и улучшить сегментацию границы опухоли. Для решения задачи предсказания выживаемости предлагается регрессионная модель СохРН относительной опасности, использующая комбинацию клинических признаков, а также признаков, полученных при глубоком обучении на ПЭТ/КТ-изображениях.

Для задачи сегментации была использована трилинейная интерполяция ПЭТ и КТ-изображений. Интенсивность ПЭТ была нормализована с помощью Z-score, а интенсивность КТ приведена к $[-1,1]$. Данные для предсказания выживаемости были обработаны с учетом пропущенных значений.

Архитектура предложенной сети NormResSE-UNet3+:

- На вход подается тензор, размерности $2 \times 144 \times 144 \times 144$, состоящий из конкатенации ПЭТ и КТ изображений.
- Энкодер состоит из residual squeeze-and-excitation блоков, первый блок из которых содержит 24 фильтра. Размерность выхода энкодера - $384 \times 3 \times 3 \times 3$
- Путь декодирования состоит из полномасштабных соединений и модуля, содержащего правильную разметку изображений (ground truth).
- У декодера одноканальный выход размерности $1 \times 144 \times 144 \times 144$



Было обучено несколько моделей нейронных сетей для задачи сегментации опухолей головы и шеи. Предложенная регрессионная модель SoxPH, а также модель NormResSE-UNet3+ показали лучший результат по сравнению с другими предшествующими моделями.

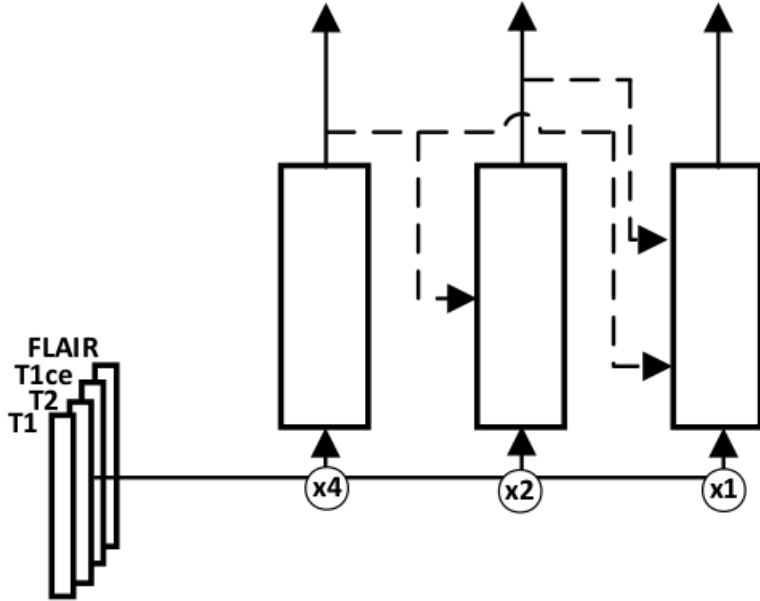
Cross-validation fold	NormResSE-UNet3+		NormResSE-UNet3+ + CRF	
	DSC	HD	DSC	HD95
Fold 1	0.792	3.18	0.822	3.11
Fold 2	0.693	3.43	0.702	3.41
Fold 3	0.728	3.32	0.749	3.29
Fold 4	0.736	3.31	0.738	3.30
Fold 5	0.742	3.29	0.756	3.28
Ensemble	0.738	3.30	0.753	3.28

Cascaded Unet

Точная сегментация и реконструкция медицинских 3D изображений способны дать больше необходимой информации о прогрессировании заболевания и позволяют терапевту спланировать успешный курс лечения для больного. В данной работе [5] авторы представляют каскадный вариант популярной сети UNet [4], который итеративно улучшает результаты сегментации, полученные на предыдущих шагах.

Предложенный метод может быть представлен как цепь классификаторов C_i , одинаковой топологии F , у каждого из которых свой собственный набор параметров W_i для оптимизации в течение обучения. Результат вычисления i -го шага представляется следующим образом: $Y_i = F(X_i, Y_{i-1}, Y_{i-2}, W_i)$. Каждый из базовых блоков C_i - это сеть архитектуры UNet, измененная для задачи сегментации глиом. В сравнении со стандартной архитектурой UNet,

в предложенной модели используется несколько энкодеров, которые отдельно обрабатывают входные данные. Также, предложен метод объединения их выхода: в UNet i -й выход декодера зависит от выхода соответствующего энкодера и выхода предыдущего декодера - $d_i^t = f(e_i^t, d_{i-1}^t)$. Раскрывая первую свертку f , получаем - $d_i^t = g(W_{i,e}^t e_i^t + W_i^t d_{i-1}^t)$. Далее предлагается объединить контекст, полученный на более низких слоях, добавляя соответствующий выход y^t , поэтому $d_i^t = g(W_{i,e}^t e_i^t + W_{i,d}^t d_{i-1}^t + W_{i,y}^t y^{t-i})$.



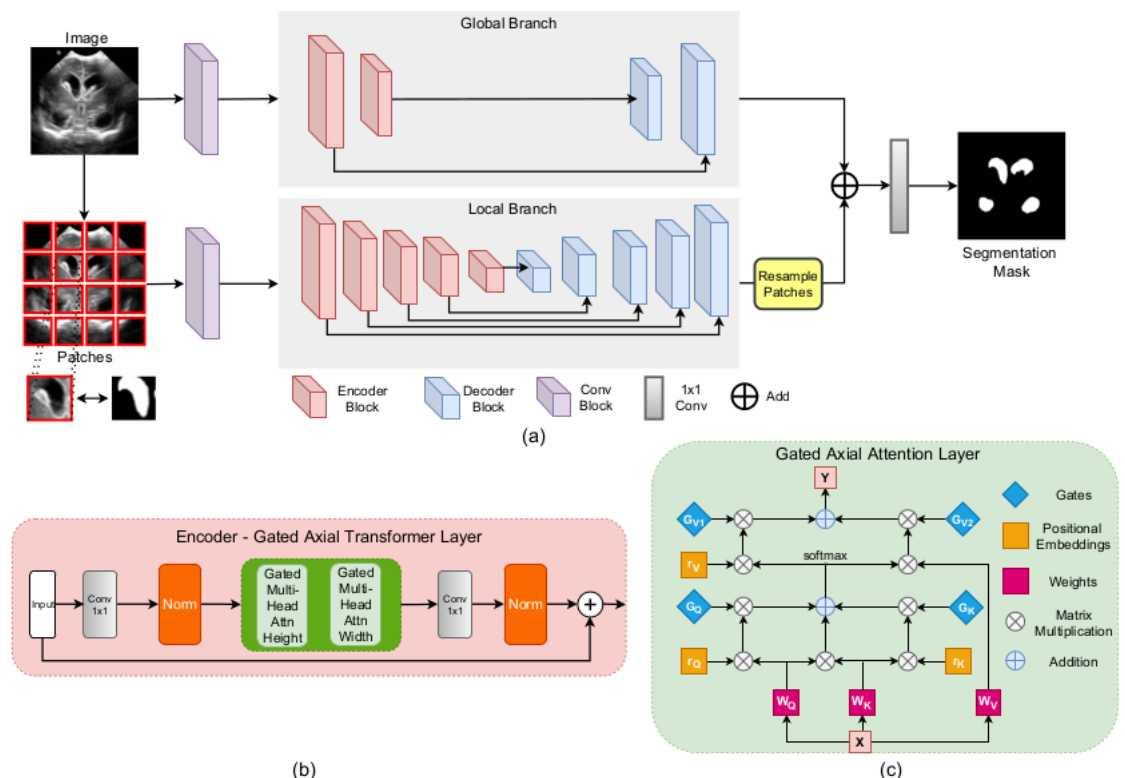
В качестве датсета использовался набор МРТ-изображений BraTS2018 [6]. Результат сегментации оценивался по метрике Dice, отдельно вычисленной для следующих частей опухоли: WT (whole tumor) - вся опухоль, ET (enhancing tumor) - усиливающаяся часть опухоли и TC (tumor core) - ядро опухоли. Предложенный алгоритм автоматической сегментации опухолей головного мозга по МРТ-зображениям, который решает также проблему мультимодального входа и показывает хорошие результаты по сравнению с моделью UNet.

Medical Transformer

Сверточные нейронные сети сравнительно плохо понимают зависимости между признаками, которые находятся на дальнем расстоянии друг от друга в изображениях. Недавно предложенные архитектуры сетей, основанные на трансформерах [7], используют механизм самовнимания [8] для шифрования дальних зависимостей и выявляют наиболее заметные представле-

ния. Эти наблюдения мотивировали авторов статьи исследовать решения, в основе которых лежат трансформеры и изучить возможность использования архитектур нейронных сетей с трансформерами в задачах медицинской сегментации. [9]

В данной работе предлагается закрытый, чувствительный к расположению axial attention механизм, который хорошо показывает себя на малых наборах данных. Также, вводится эффективная методология обучения Local-Global (LoGo) для трансформеров и Medical-Transformer (MedT) - метод, построенный на основе двух вышеперечисленных предложенных концепций, разработанный специально для сегментации медицинских изображений, который успешно улучшает производительность по сравнению со сверточными нейронными сетями и сетями чисто attention архитектуры на трех разных датасетах. Предложенные методы превосходят существующие в задаче сегментации медицинских изображений не требуя при этом большого набора тренировочных данных.



Local-Global training:

Чтобы улучшить общее понимание изображения, предлагается использовать два ответвления сети - глобальная ветвь, которая работает с изображением оригинальной размерности и локальная ветвь для работы с патчами.

В локальной ветви создается 16 патчей размера $I/4 \times I/4$, каждый патч пропускается через сеть и выходные карты признаков resampled, основываясь на их расположении, чтобы получить итоговые карты признаков. К итоговым картам двух ветвей применяется свертка 1×1 и на выходе получается маска сегментации. Такая стратегия улучшает производительность, так как глобальная ветвь фокусируется на высокоуровневой информации, а локальная лучше определяет детали.

AFTer-UNet

Последние достижения в моделях, основанных на трансформерах [7] притянули внимание исследователей для изучения данных техник в сегментации медицинских изображений. Особенно часто трансформеры используются совместно с U-Net подобными моделями (и их производными), которые показали хороший результат в обработке двумерных и трехмерных изображений. В существующих методах, основанных на двумерных изображениях, сверточные слои напрямую заменяются трансформерами, либо трансформер применяется в качестве дополнительного промежуточного энкодера между энкодером и декодером U-Net. Однако, при таком подходе пространственная информация используется не полностью, сеть видит только срезы трехмерного изображения, не учитывая связи в целом трехмерном изображении.

Для того, чтобы трансформер мог находить связи между признаками, находящимися на далеком расстоянии в трехмерных медицинских изображениях, в данной работе [10] предлагается архитектура Axial Fusion Transformer UNet (AFTer-UNet). AFTer-UNet следует архитектуре U-Net [4], которая состоит из двумерного сверточного энкодера и декодера, но между ними авторы разместили промежуточный axial fusion transformer энкодер, для того, чтобы связать контекстную информацию из соседних срезов. Промежуточный энкодер снижает вычислительную сложность тем, что сначала отдельно вычисляется фокус (attention) вдоль осей и внутри одного среза, а потом полученная информация объединяется для составления финальной карты сегментации.

Vit-V-Net

В данной работе [11] исследуется применение ViT [7] в объемных медицинских изображениях. Авторы предлагают ViT-V-Net [12], которая воплощает в себе гибридную архитектуру „сверточная нейронная сеть-трансформер“ (ConvNet-Transformer) для применения self-supervised метода в исследовании трехмерных медицинских изображений.

В предложенном методе ViT была применена к высокоуровневым признакам изображений, что требовало от сети выявить зависимости между точками, находящимися на дальнем расстоянии. Наивное применение ViT к полномасштабным изображениям приводит к увеличению вычислительной сложности. Поэтому, изображения сначала были закодированы с помощью нескольких сверточных слоев и слоев max-pooling для получения объектов, содержащих высокоуровневые признаки. Далее, в ViT, высокоуровневые признаки делятся на патчи, а затем патчи отображаются в скрытое пространство с помощью обучаемой линейной проекции (например, patch embedding). Затем, результирующие патчи подаются в энкодер трансформера, а полученный выход декодируется V-Net подобным декодером.

CaraNet

Только малая часть исследований учитывает размер интересующих объектов на изображении и поэтому многие модели показывают плохой результат при сегментации объектов малого размера, что сильно влияет на диагностику заболевания. В данной работе [13] предлагается нейросетевая модель Context Axial Reserve Attention Network (CaraNet), которая способна улучшить результаты сегментации малых объектов по сравнению с уже существующими моделями.

В архитектуре CaraNet используется параллельный частичный декодер (parallel partial decoder) для генерации высокоуровневой семантической карты и набор операций (Context and Axial Reverse Attention) для идентификации глобальных и локальных признаков.

Модули CaraNet:

- *Parallel partial decoder.* Эксперименты показали, что низкоуровневые признаки вычислительно более сложны и вносят меньший вклад в улуч-

шение результатов сегментации. Поэтому, авторы используют параллельный частичный декодер $p_d(\cdot)$ для извлечения высокоуровневых признаков $PD = p_d(f_3, f_4, f_5)$ и получения глобальной карты S_g из частичного декодера.

- *Context module.* Чтобы получить контекстную информацию из высокоуровневых признаков, применяется модуль CFP (Channel-wise Feature Pyramid) со степенью растяжения (dilation rate) $d = 8$. После контекстного модуля можно получить многомасштабные высокоуровневые признаки $\{f'_3, f'_4, f'_5\}$.
- *Axial reverse attention.* Данный модуль состоит из двух частей: маршрут по оси (axial attention route) и обратный маршрут (reverse attention route). Глобальная карта S_g может поймать только приблизительное расположение тканей без структурных деталей, поэтому структурированный регион тканей постепенно добывается стиранием переднего плана объекта с помощью операции reverse attention: $R_i = 1 - \text{Sigmoid}(S_i)$. По другому маршруту применяется axial attention. Здесь сеть может извлекать глобальные зависимости и локальное представление совершая вычисления по горизонтальной и вертикальной оси.

По сегментации полипов на основе пяти датасетов, предложенная модель CaraNet не только превосходит сравниваемые модели по общей производительности, но и на примерах с полипами малых размеров. Для дальнейшей оценки эффективности сегментации малых объектов с помощью CaraNet был проведен еще один эксперимент, уже с участием опухолей ГМ из датасета BraTS 2018 [6]. CaraNet была сравнена с PraNet [14] и показала лучший результат особенно в случаях с очень малыми объектами.

3D Self-Supervised Methods for Medical Imaging

В данной работе [15] предлагаются трехмерные варианты self-supervised методов, которые облегчают обучение нейронной сети на признаках по немаркированным трехмерным изображениям, что приводит к снижению затрат на экспертную аннотацию. Рассмотрены 5 алгоритмов и проведен сравнительный анализ на трехмерных медицинских изображениях (МРТ, КТ). Выбор

алгоритмов обусловлен их успешным применением в двумерном случае и тем, что ни один из них не был расширен до трехмерного на момент выхода статьи.

Авторы предлагают 5 алгоритмов, которые целиком используют пространственную информацию 3D-изображения. В каждом методе используется энкодер g_{enc} , который может быть дообучен под различные задачи.

- **3D Contrastive Predictive Coding (3D-CPC)**

Следуя идее, предложенной в двумерном случае [16], этот метод предсказывает скрытое пространство для следующих (смежных) образцов. Предложенный CPC определяет задачу, обрезая одинаковые по размеру и перекрывающиеся участки каждого сканирования. Далее, энкодер g_{enc} сопоставляет каждый входной патч $x_{i,j,k}$ его скрытому представлению $z_{i,j,k} = g_{enc}(x_{i,j,k})$. Затем, следующая модель, называемая контекстной сетью g_{cxt} суммирует скрытые вектора патчей контекста $x_{i,j,k}$ и составляет свой контекстный вектор $c_{i,j,k} = g_{cxt}(\{z_{u,v,w}\})$, где $\{z\}$ - это множество скрытых векторов. Наконец, так как $c_{i,j,k}$ захватывает высокоуровневый контент из контекста, который отвечает $x_{i,j,k}$, это позволяет предсказать скрытые представления следующих(смежных) патчей $z_{i+l,j,k}$, где $l \geq 0$. Стоит отметить, что в предложенном 3D-CPC в качестве g_{enc} и g_{cxt} могут использоваться сети любой архитектуры.

- **Relative 3D patch location (3D-RPL)**

В этой задаче пространственный контекст в изображениях используется как богатый источник для семантического представления данных. В предложенной 3D версии из каждого входного 3D изображения выбирается сетка из N неперекрывающихся участков $\{x_i\}_{i \in \{1, \dots, N\}}$ случайного расположения. Далее, центральный патч x_c используется как ссылка, а очередной патч x_q выбирается из окружающих $N - 1$ патчей. Далее, расположение x_q относительно x_c выбирается как положительная метка y_q . Таким образом, задача сводится к $N - 1$ -классовой классификации, где расположения оставшихся патчей используются как негативные метки.

- **3D Jigsaw puzzle Solving (3D-Jig)**

Получение мозаичной сетки из входного изображения может рассматриваться как расширение вышеприведенной задачи RPL, основанной на патчах. Пазлы формируются путем выбора $n \times n \times n$ сетки из 3D пат-

чей, далее эти патчи перемешиваются следуя произвольной перестановке из множества предопределенных перестановок с индексом $y_p \in \{1, \dots, P\}$, где P - размерность множества перестановок, выбранного из $n^3!$ всевозможных перестановок. Таким образом, задача сводится к P -классовой классификации - модель тренируется просто запомнить индекс p примененной перестановки.

- **3D Rotation prediction (3D-Rot)**

В данной задаче модель должна предсказать угол, на который повернуто изображение. Входное изображение поворачивается случайным образом на угол $r \in \{1, \dots, R\}$. Поворот изображения на угол в 0° вдоль трех осей произведет три идентичных версии исходного изображения, поэтому рассматриваются только 10 возможных поворотов из 12. В таких условиях задача сводится к 10-классовой классификации.

- **3D Exemplar networks (3D-Exe)**

Для получения supervised-меток метод опирается на аугментацию изображений. Здесь для тренировочного набора данных определяется множество трансформаций изображения, а новый суррогатный класс создается с помощью трансформации тренировочного примера. Задача является обычной задачей классификации с кросс-энтропийной функцией потерь. Однако, с увеличением датасета и количества классов задача становится более вычислительно сложной, поэтому в предложенной 3D версии внедрен механизм, который опирается на тройную функцию потерь.

Предложенные методы были опробованы в различных медицинских задачах и показали следующие результаты:

- 1) Сегментация опухолей мозга: все предложенные методы преодолевают бейзлайны, так же, как и двумерные версии этих методов. Результаты, полученные в данной задаче показывают наличие обобщающей способности у всех предложенных методов.
- 2) Сегментация опухолей поджелудочной железы: результаты, полученные с помощью предложенных методов преодолевают бейзлайны для поставленной задачи. Также, предложенные методы показывают достаточно быструю сходимость.

Полученные результаты являются конкурентноспособными, а разработанные методы могут применяться в дальнейших исследованиях.

Список литературы

1. Первичные опухоли центральной нервной системы // Клинические рекомендации. — 2019. — т. 58.
2. Молекулярная неврология, часть III Опухоли головного мозга, онкогены и антионкогены / В. Горбунова [и др.]. — Интермедика, 2004.
3. Multimodal PET/CT Tumour Segmentation and Prediction of Progression-Free Survival using a Full-Scale UNet with Attention / B. Emmanuelle [и др.] // arXiv:2111.03848v1. — 2021. — т. 13, eess.IV.
4. *Ronneberger O., Fischer P., T. B.* U-net: Convolutional networks for biomedical image segmentation. // International Conference on Medical image computing and computer-assisted intervention. — 2015. — т. 7.
5. *Lachinov D., Vasiliev E., Turlapov V.* Glioma Segmentation with Cascaded Unet // arXiv:1810.04008. — 2018. — т. 9, cs.CV.
6. Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features / B. S. [и др.] // Nature Scientific Data. — 2017. — 10.1038/sdata.2017.117.
7. An image is worth 16x16 words: Transformers for image recognition at scale / A. Dosovitskiy [и др.] // arXiv. — 2020. — ArXiv:2010.11929.
8. Attention Is All You Need / A. Vaswani [и др.] // arxiv. — 2017. — т. 15, arXiv:1706.03762.
9. Medical Transformer: Gated Axial-Attention for Medical Image Segmentation / J. M. J. Valanarasu [и др.]. — 2021.
10. AFTer-UNet: Axial Fusion Transformer UNet for Medical Image Segmentation / X. Yan [и др.] // arxiv. — 2021. — т. 11, arXiv:2110.10403.
11. ViT-V-Net: Vision Transformer for Unsupervised Volumetric Medical Image Registration / J. Chen [и др.] // arxiv. — 2021. — т. 9, arXiv:2104.06468.

12. *Milletari F., Navab N., Ahmadi S.-A.* V-net: Fully convolutional neural networks for volumetric medical image segmentation // Fourth international conference on 3D vision (3DV). — 2016. — т. 6.
13. CaraNet: Context Axial Reverse Attention Network for Segmentation of Small Medical Objects / A. Lou [и др.] // arxiv. — 2021. — т. 12, arXiv:2108.07368.
14. Pranet: Parallel reverse attention network for polyp segmentation / D. P. Fan [и др.] // International Conference on Medical Image Computing and Computer-Assisted Intervention. — 2020. — т. 10.
15. 3D Self-Supervised Methods for Medical Imaging / A. Taleb [и др.] // arXiv. — 2020. — т. 19, arXiv:2006.03829.
16. *Oord A. van den, Li Y., Vinyals O.* Representation learning with contrastive predictive coding // arXiv. — 2018. — Abs/1807.03748.