

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ

Направление 01.03.02 Прикладная математика и информатика
Кафедра технологии программирования

ОТЧЁТ
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ
по теме:
ИССЛЕДОВАНИЕ НЕЙРОСЕТЕВЫХ МЕТОДОВ
СЕГМЕНТАЦИИ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

Научный руководитель _____ А.Г. Головкина
кандидат физико-математических наук, доцент

Работу выполнила _____ А.А. Иванова
группа 20.Б08-пу

Санкт-Петербург 2022

СОДЕРЖАНИЕ

Введение	3
1 Постановка задачи	4
1.1 Данные соревнования	4
2 Возможные подходы к решению задачи	6
2.1 2D подход	6
2.2 3D подход	7
2.3 2.5D подходы к сегментации	8
2.3.1 Включения межсрезовой информации	9
2.3.2 Смешение 2D/3D признаков	10
Заключение	11
Список использованных источников	12

ВВЕДЕНИЕ

Анализ медицинских изображений играет важную роль в оказании помощи практикующим врачам в различных экспертных исследованиях, клинической диагностике и выборе дальнейшего оптимального плана лечения. Одними из основных задач анализа медицинских изображений являются задачи сегментации и классификации различных биологических структур на изображении, таких как ткани и органы.

В последнее время глубокие сверточные нейронные сети (CNN - convolutional neural network [1]) достигли больших успехов в этих задачах. Однако, в отличие от естественных изображений, большинство медицинских изображений, таких как МРТ и КТ, представляют собой трехмерные объемные данные. Из-за большей размерности входных данных, глубокие нейронные сети имеют больше параметров, а значит требуют больше вычислительных мощностей, количества данных для обучения и более длительного времени обучения.

Целью данной работы является исследование нейросетевых методов сегментации медицинских изображений и дальнейшая разработка метода выявления переломов на КТ шейного отдела позвоночника (шеи) как на уровне отдельных позвонков, так и у всего больного. Для этого нужно:

- провести изучение предметной области и анализ существующих решений поставленной задачи на основе методов глубокого обучения;
- провести предобработку данных;
- разработать архитектуру, позволяющую эффективно решать поставленную задачу;
- произвести обучение реализованной модели;
- провести вычислительный эксперимент с целью оценки результатов, полученных на основе предлагаемых методов и алгоритмов.

1 Постановка задачи

В данной научно-исследовательской работе будет рассмотрено возможное решение для соревнования RSNA 2022 Cervical Spine Fracture Detection [2] на платформе Kaggle [3].

Задача обнаружения перелома шейного отдела позвоночника представляет собой последовательность из нескольких задач: во-первых, сегментации каждого позвонка, во-вторых, определение наличия или отсутствия перелома, а также класса каждого позвонка, то есть задачи классификации.

Наиболее частой локализацией перелома позвоночника является шейный отдел позвоночника. Визуальная диагностика переломов позвоночника у взрослых в настоящее время почти всегда выполняется с помощью компьютерной томографии (КТ) вместо рентгеновских снимков. Быстрое обнаружение и определение местоположения любых переломов позвонков имеет важное значение для предотвращения неврологического ухудшения состояния пациента и паралича после травмы.

1.1 Данные соревнования

Данные соревнования включают в себя примерно 3000 исследований КТ. Специалисты по радиологии позвоночника из ASNR и ASSR предоставили экспертные аннотации на уровне изображений для этих исследований, чтобы указать наличие, уровень позвонков и местоположение любых переломов шейного отдела позвоночника.

Каждое изображение находится в формате файла DICOM. Стандарт DICOM — Digital Imaging and Communications in Medicine [4] — используется для обмена снимками и информацией, он состоит из нескольких слоёв, что позволяет отправителям и получателям снимков обмениваться информацией об анализируемых изображениях.

Нам необходимо предсказать вероятность перелома для каждого из семи шейных позвонков, обозначенных C1, C2, C3, C4, C5, C6

и С7, а также общую вероятность любых переломов шейного отдела позвоночника. Для некоторых исследований КТ известны маски для сегментации и ограничивающие рамки (bounding boxes) областей, где произошел перелом.

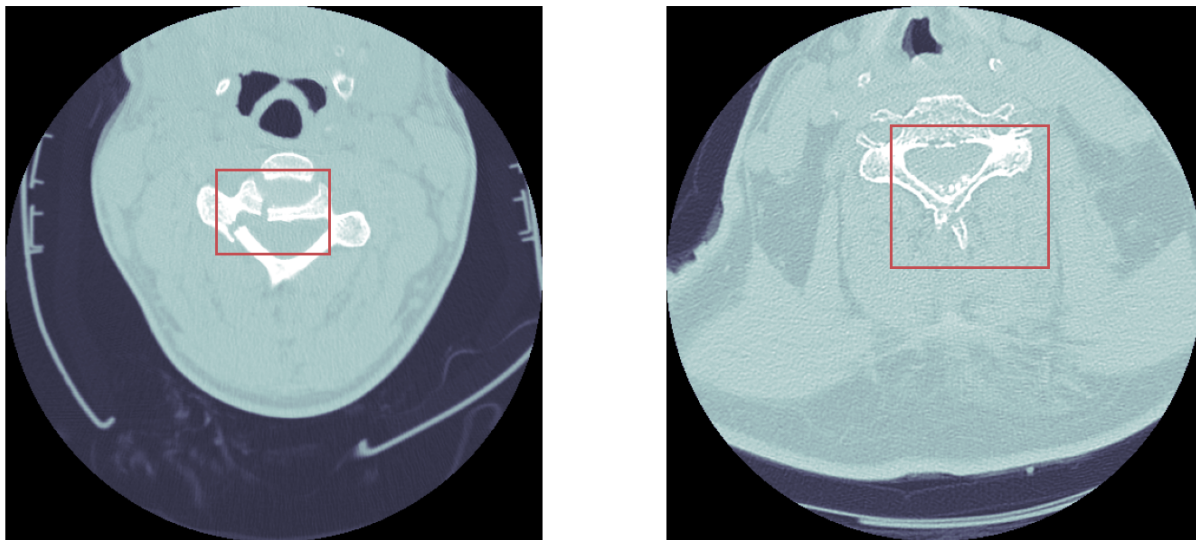


Рисунок 1.1 — Визуализация ограничивающих рамок (bounding boxes) перелома

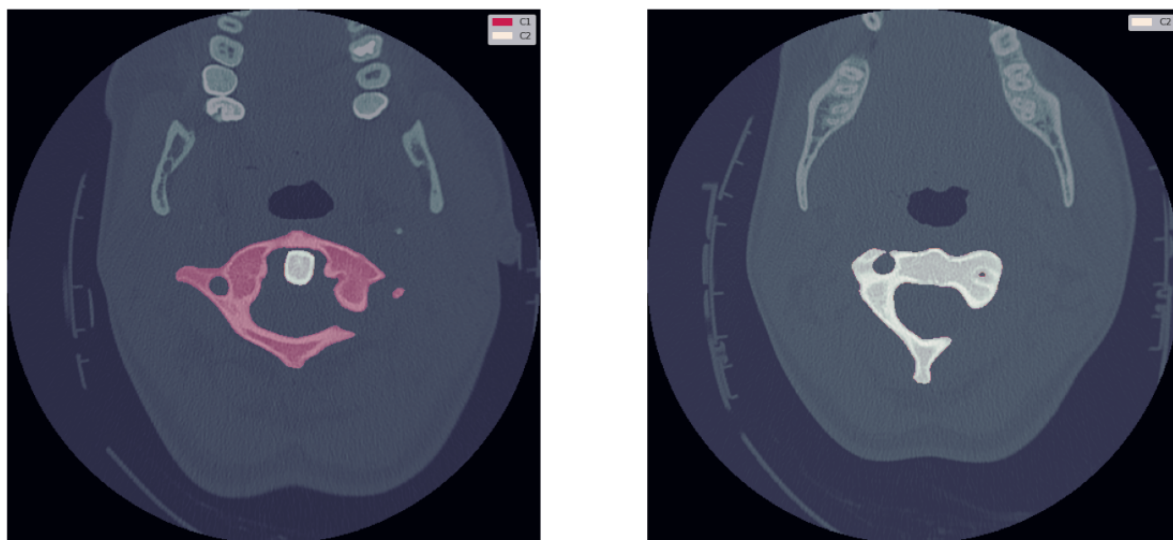


Рисунок 1.2 — Визуализация масок сегментации позвонков

2 Возможные подходы к решению задачи

Наиболее перспективным направлением в медицине считается применение сверточных нейронных сетей для обработки изображений. [5, 6]

2.1 2D подход

Наиболее популярной архитектурой для сегментации медицинских данных является U-Net [7], на основе которой в последствии было разработано множество других вариаций, например

Основой данной архитектуры является наличие двух частей: сужаю-

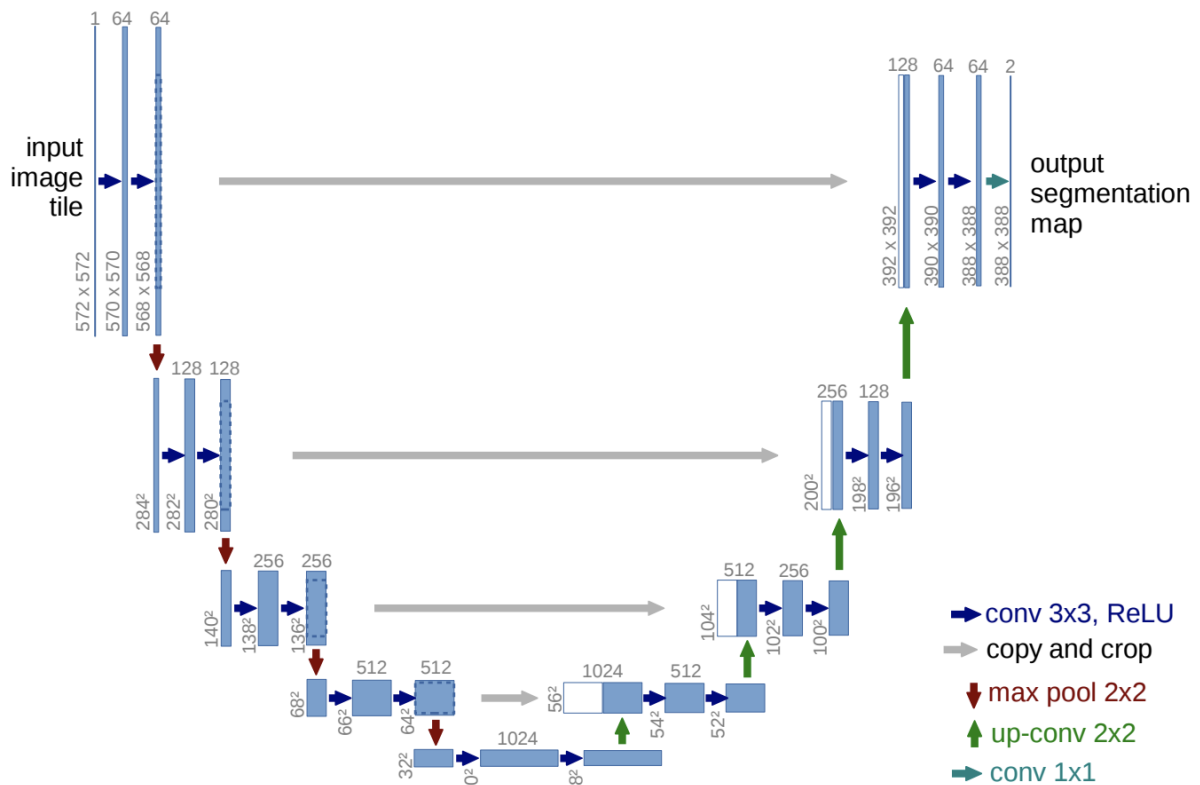


Рисунок 2.1 — Архитектура 2D UNet

щейся (энкодер) и расширяющейся (декодер). Энкодер представляет собой типичную архитектуру сверточной сети и состоит из сверточных слоев (convolution), за каждой из которых следует операция активации (ReLU), а затем операция max-pooling'a. Декодер состоит из шагов обратной свертки (deconvolution), которая уменьшает количество каналов признаков, после чего следует конкатенация с обре-

занной картой признаков от соответствующей части энкодера (skip connection) и основной блок. Обрезка нужна из-за того, что мы теряем пограничные пиксели в каждой свёртке.

Причина, по которой U-Net подходит для сегментации медицинских изображений, заключается в том, что ее структура может одновременно объединять низкоуровневую и высокоуровневую информацию. Информация низкого уровня помогает повысить точность. Информация высокого уровня помогает извлекать сложные признаки изображения.

Главным недостатком 2D подхода для сегментации КТ снимков является то, что теряется важная пространственная информация, которая могла бы быть использована для повышения точности сегментации.

2.2 3D подход

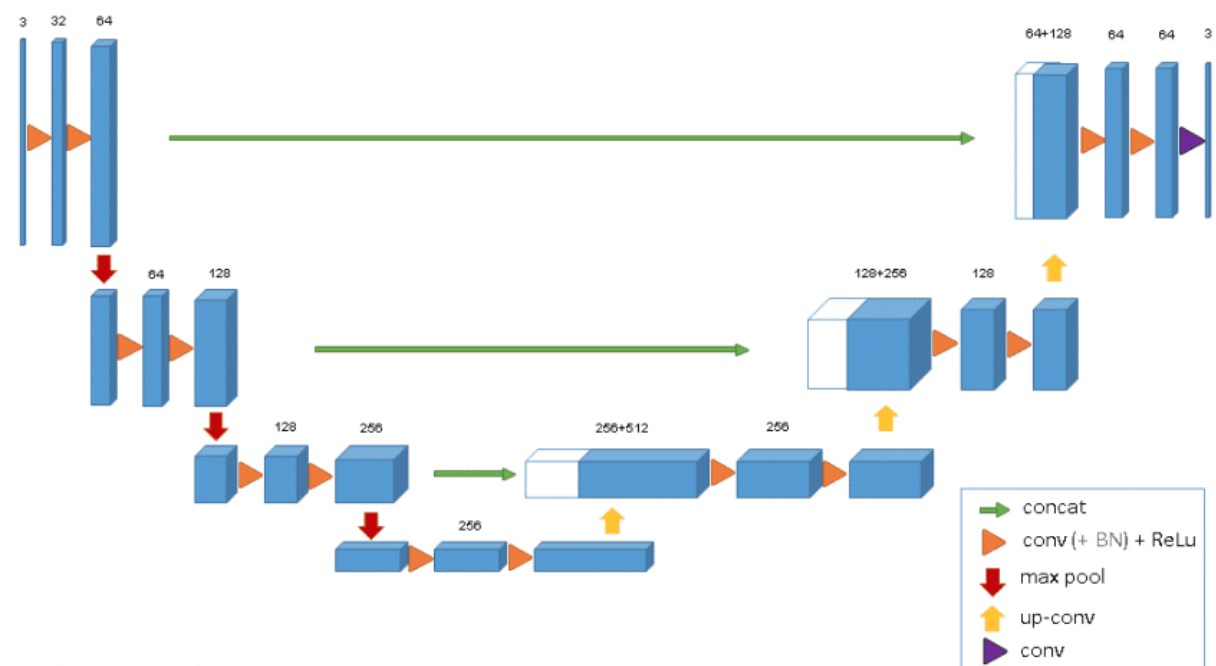


Рисунок 2.2 — Архитектура 3D UNet

Для сегментации объемных медицинских изображений 3D CNN могут использовать более богатую контекстную информацию и улучшить производительность по сравнению с 2D CNN, но в то же

время это ставит более высокие требования к памяти и увеличивает время обучения моделей. [8]

Например, естественным обобщением 2D UNet стала архитектура 3D UNet [9], повторяющая исходную архитектуру, за исключением того, что операции свертки и пулинга стали проходить в трехмерном пространстве. Слои свертки выполняют фильтрацию с помощью 3D-ядер, а слои 3D-пулинга субдискретизируют данные во всех трех измерениях, чтобы сжать размер пространства признаков. Таким образом, объемные данные анализируются как кубические фрагменты, что позволяет изучать пространственные особенности во всех трех измерениях.

В связи с вычислительной сложностью обучения таких громоздких моделей и необходимостью сбора больших датасетов, встает вопрос о разработке менее ресурсоемких подходов, которые могли бы обучаться с меньшим количеством параметров и меньшим количеством аннотированных медицинских данных для обучения, но были бы сравнимы по точности результата с трехмерными архитектурами.

2.3 2.5D подходы к сегментации

Чтобы решить вышеописанную проблему, во многих исследованиях предлагается использовать сегментацию так называемыми 2.5D методами, которые объединяют частичное использование объемной пространственной информации и 2D CNN для повышения точности при одновременном снижении вычислительных затрат. [10] Этого можно добиться несколькими способами, например с помощью включения межсрезовой информации (incorporating inter-slice information) или смешения 2D/3D признаков (fusing 2D/3D features).

2.3.1 Включения межсрезовой информации

При помощи включения межсрезовой информации сеть может использовать не только информацию об имеющемся срезе, но и информацию о соседних срезах, что позволяет частично использовать пространственную информацию, избегая при этом тяжелой нагрузки трехмерных вычислений. Наиболее широко используемый метод основан на идее 2D сегментации изображения, он вводит последовательность срезов в качестве многоканального ввода для сегментации среднего среза.

Возможные решения для более точных результатов 2D-сегментации может обеспечить применение рекуррентных нейронных сетей (RNN). Таким образом, 2D-срезы 3D-объема рассматриваются как последовательность временных рядов для выделения контекста и признаков между срезами.

Кроме использования RNN, другое направление исследований сосредоточено на использовании механизма внимания (attention mechanism). Поскольку информация каждого среза пространственно коррелирует с его верхним и нижним срезами, из-за пространственной непрерывности информация о соседних срезах может использоваться для руководства процедурой сегментации путем выделения наиболее заметных областей с использованием механизма внимания.

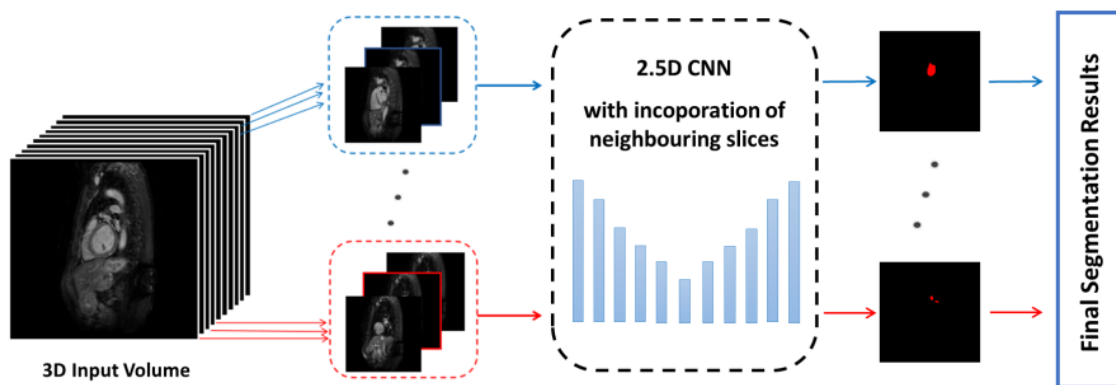


Рисунок 2.3 — Включения межсрезовой информации в 2.5D CNN

2.3.2 Смешение 2D/3D признаков

Смешение 2D/3D признаков представляет собой методы, использующие одновременно 3D и 2D блоки для извлечения 3D и 2D признаков соответственно. Хотя эти методы по-прежнему используют трехмерные свертки для извлечения пространственной информации, общая стоимость вычислений снижается по сравнению с обучением чистых трехмерных CNN.

Один из основных методов данной категории использует 2D-результаты сегментации для предоставления информации о форме для 3D-CNN, которая затем использует пространственную информацию на основе этих 2D-выходов и исходного объема. В конце концов, результаты 2D и 3D объединяются, чтобы получить окончательные результаты сегментации.

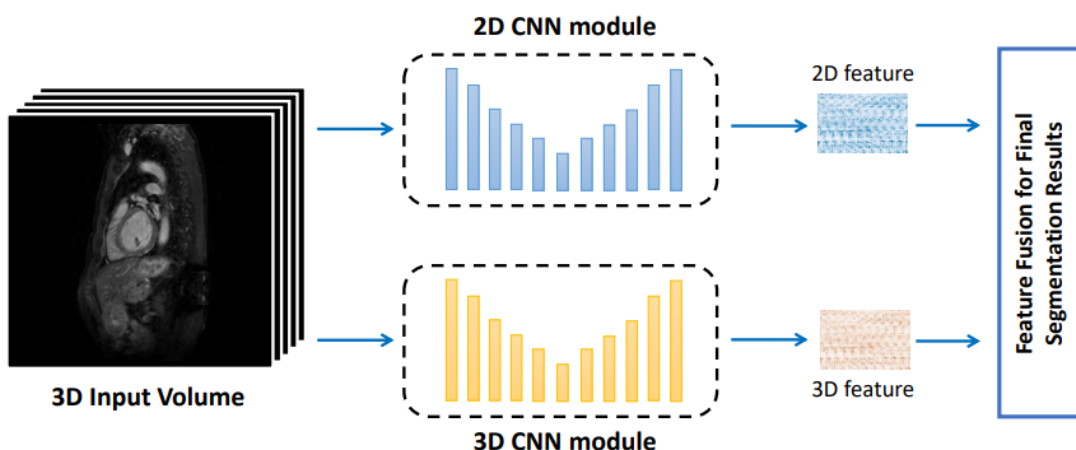


Рисунок 2.4 — Смешение 2D/3D признаков в 2.5D CNN для уточнения итоговой сегментации

ЗАКЛЮЧЕНИЕ

В результате проделанной работы был проанализирован ряд статей, позволяющих получить общее представление о задаче сегментации медицинских многомерных изображений на основе глубокого обучения.

В дальнейшем планируется проведение дальнейшего изучения темы, предобработка имеющихся данных соревнования, а затем разработка архитектуры на основе уже имеющихся методов и обучение сети в условиях ограниченной вычислительной мощности.

Несмотря на полученные результаты, вопрос того, какую именно архитектуру сети для сегментации выбрать, требует дальнейшего исследования.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Gradient-based learning applied to document recognition. / Lecun Y., Bottou L., Bengio Y., and Haffner P. // Proceedings of the IEEE. — 1998. — Vol. 86, no. 11. — P. 2278–2324.
2. RSNA 2022 Cervical Spine Fracture Detection. — Access mode: <https://www.kaggle.com/competitions/rsna-2022-cervical-spine-fracture-detection/overview> (online; accessed: 19.12.2022).
3. Kaggle. — Access mode: <https://www.kaggle.com> (online; accessed: 18.12.2022).
4. DICOM standart official page. — Access mode: <https://www.dicomstandard.org/current> (online; accessed: 19.12.2022).
5. Yousef R. Gupta G. Yousef N. A holistic overview of deep learning approach in medical imaging. // Proceedings of the IEEE. — 2022. — Vol. 28. — P. 881–914.
6. A Review of Deep-Learning-Based Medical Image Segmentation Methods / Liu Xiangbin, Song Liping, Liu Shuai, and Zhang Yudong // Sustainability. — 2021. — Vol. 13, no. 3. — Access mode: <https://www.mdpi.com/2071-1050/13/3/1224>.
7. Ronneberger Olaf, Fischer Philipp, Brox Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation. — 2015. — Access mode: <https://arxiv.org/abs/1505.04597>.
8. Medical Image Segmentation with 3D Convolutional Neural Networks: A Survey. — 2021. — Access mode: <https://arxiv.org/abs/2108.08467>.
9. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. — 2016. — Access mode: <https://arxiv.org/abs/1606.06650>.

10. Bridging 2D and 3D Segmentation Networks for Computation Efficient Volumetric Medical Image Segmentation: An Empirical Study of 2.5D Solutions. — 2020. — Access mode: <https://arxiv.org/abs/2010.06163>.