

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Петрова Олеся Геннадьевна

Выпускная квалификационная работа

**Сегментация МРТ-изображений с использованием
нейронных сетей**

Уровень образования: магистратура

Направление 01.04.02 «Прикладная математика и информатика»

Основная образовательная программа: ВМ.5691 «Прикладная математика и информатика в задачах медицинской диагностики»

Научный руководитель:

д-р. физ.-мат. наук, профессор

кафедра теории управления

Котина Елена Дмитриевна

Рецензент

к.ф.-м.н., доцент

кафедра медицинской биофизики СПбГПМУ

Елизарова Марина Владиславовна

Санкт-Петербург

2021

Содержание

Введение.....	3
Постановка задачи.....	5
Обзор литературы.....	6
Исходные данные	11
Глава 1. Применение методов обработки изображений	13
1.1. Матрица смежности уровней серого	13
1.2. Геометрический фильтр Кримминса.....	15
1.3. Фильтр «Нерезкое маскирование»	16
1.4. Фильтры Габор.....	17
Глава 2. Архитектуры нейронных сетей для сегментации изображений	19
2.1. U-Net	19
2.2. Residual U-Net	20
2.3. ResUnet++	22
2.4. UNet++	23
Глава 3. Разработка и реализация алгоритма автоматической сегментации ..	24
3.1. Предварительная обработка изображений	25
3.2. Сегментация изображений	29
3.2.1. Обучение без дополнительной предварительной обработки.	30
3.2.2. Обучение с применением дополнительной предварительной обработки	34
Выводы	36
Заключение	37
Список литературы	38

Введение

Обработка медицинских изображений – трудоемкий процесс, требующий предельной внимательности и профессионализма от специалиста, а также определенных временных затрат. Во многих случаях быстрая и своевременная диагностика может в значительной мере облегчить и ускорить процесс выздоровления пациента.

Анализ медицинских изображений используется как основной метод диагностики многих заболеваний позвоночника. Как правило, эти заболевания сопровождаются болью в спине, которая мешает человеку нормально жить. В таком случае от скорости диагностики зависит не только то, как быстро человек получит необходимое лечение, но и то, насколько длительный курс реабилитации будет необходимо пройти.

Для исследования состояния позвоночника используется несколько способов:

1. Традиционная рентгенография позвоночника. Используется в качестве первоначального этапа исследования состояния позвоночного столба и позволяет выявить только явные проблемы – значительные переломы и трещины, а также оценить общее состояние костей и позвонков. Для уточнения картины заболевания или в случае, когда симптомы сохраняются после проведенного лечения, назначаются дополнительные исследования (КТ или МРТ).

2. Компьютерная томография (КТ) – наиболее эффективный метод для исследования состояния костных структур, диагностики сложных травм позвоночника и кровотечений.

3. Магнитно-резонансная томография (МРТ) – позволяет получить полное представление о состоянии и строении мягких тканей. Используется для обнаружения опухолей или воспалительных изменений в суставах, межпозвоночных дисках, спинном мозге, а также для обнаружения деформаций при остеохондрозе, грыже или протрузии.

Актуальность исследования: По данным ВОЗ примерно 1.71 миллиарда человек по всему миру страдают от нарушений и болезней костно-мышечной системы, из них 568 миллионов приходится на люмбаго (острая боль в поясничном отделе позвоночника, независимо от причины возникновения) [1]. Зачастую длительная боль в пояснице оказывает значительное влияние на психоэмоциональное состояние человека, а также на другие сферы его жизни. Кроме того, боль в пояснице считается одной из самых частых причин инвалидности во всем мире [2].

К основным факторам риска развития люмбаго относятся: возраст старше 30 лет, плохая физическая форма, сидячая работа или работа, связанная с поднятием тяжестей, стрессы и нездоровый образ жизни. По прогнозам, количество людей, страдающих от болей в пояснице, будет возрастать, в странах с низким и средним уровнем жизни темпы роста будут наиболее высоки [3].

В последнее время активно идет разработка и создание средств для автоматизации обработки и анализа медицинских изображений, в том числе МРТ изображений позвоночника. Использование данных методов позволит ускорить процесс обработки изображений, кроме того, подобные программы могут обеспечить более раннюю диагностику, обнаружив физические отклонения малых размеров, которые врач еще не может заметить. Подавляющее большинство подобных разработок ведется зарубежными исследователями.

Также существуют коммерческие решения зарубежных компаний. Однако, их использование в России зачастую затруднительно из-за высокой стоимости, сложностей перевода медицинских терминов, отсутствия всех необходимых для специалистов функций и т.д.

Постановка задачи

Целью данной работы является разработка метода автоматической сегментации МРТ изображений, который позволит выделить каждый позвонок и позвоночный отросток.

В рамках поставленной цели были выделены следующие этапы:

1. **Предварительная обработка изображений.** Выбор методов обработки изображений на основе анализа литературы и их применение. Подготовка изображений к обучению.
2. **Сегментация изображений.** Исследование существующих архитектур сверточных нейронных сетей, используемых для сегментации изображений. Выбор подходящих для данной задачи.
3. **Оценка полученных результатов.** Сравнение результатов сегментации изображений для моделей с различными архитектурами и методами предварительной обработки изображений.

Обзор литературы

Подавляющее большинство предлагаемых методов сегментации изображений позвоночника основаны на применение нейронных сетей, которые являются довольно чувствительными к качеству изображений. Поэтому разработку целесообразно разделить на две части: предварительная обработка изображений и разработка нейронной сети для сегментации.

Предварительная обработка. Во время анализа существующих исследований, посвященных сегментации позвоночника, было замечено, что обычно никакой особой предварительной обработки либо не проводится, либо она не указана. По этой причине было рассмотрено применение различных методов предварительной обработки изображений в других областях, связанных с медико-биологическими данными.

Существует большое количество различных методов обработки и анализа изображений, которые принято разделять на несколько групп: статистические, геометрические, спектральные, структурные и модельные. В данной работе рассмотрены нескольких подходов.

Наиболее известным методом, который относят к *статистическим*, является метод, основанный на матрице смежности уровней серого Харалика. Примером является применение в цитологии для обнаружения раковых клеток в крови с точностью 100% и здоровых клеток с точностью 90.4% [4]. Гораздо большее распространение данный метод имеет в задачах классификации.

К *спектральным* методам относят фильтр Габора, который уже долгое время применяется в медицинской сегментации. Наиболее популярным направлением является сегментация кровеносных сосудов на изображениях сетчатки глаза [5]. В данной работе авторы подобрали оптимальные значения параметров фильтра, а также сравнили полученный результат с работами других авторов.

В источнике [6] предлагается использовать фильтры Габора в качестве ядра свертки в нейронной сети с архитектурой схожей с U-Net. Такой подход

позволяет получить больше информации о текстуре изображения. Также в данной работе проводится сравнение различных методов, которое показывает, что новая архитектура значительно лучше предложенных ранее.

В работе [7] авторы предлагают использовать фильтр Габор для автоматической оценки положения позвонков человека на изображениях в сагиттальной проекции, полученных с помощью остеоденситометрии. В статье приведены результаты эффективности, а также алгоритм фильтрации Габор в множественных направлениях.

Еще одной областью успешного применения фильтров Габор является обработка результатов коронарографии. В статье [8] описывается алгоритм автоматической сегментации коронарных артерий с использованием многокритериальной оптимизации. Также авторами приводится метод поиска наиболее оптимальных параметров для фильтра и масштабное сравнение предложенного метода с уже существующими.

Одним из направлений предварительной обработки медицинских изображений является повышение их качества, в частности снижение уровня шума. Основной целью методов данной группы является воссоздание изображения без шумов, сохраняя при этом все важные детали такие как края, текстуры и т.д. В статьях [9, 10] проведен анализ и сравнение наиболее популярных методов снижения уровня шума для различных исследований ([9] – МРТ мозга, [10] – ультразвуковая диагностика сонной артерии), а также описаны некоторые их преимущества и недостатки. Авторами работы [9] сделан вывод, что на данный момент нет такого метода, который позволит выполнить требования эффективного снижения уровня шума, однако эта область развивается и в последние годы пользуется большим интересом исследователей.

На основе анализа существующих исследований было выбрано несколько вариантов предварительной обработки изображений. Однако, так как применение методов текстурной обработки на МРТ изображениях позвоночника не было встречено, сложно сказать, какое влияние они окажут

на результат. Многие из рассмотренных методов показывают себя наилучшим образом в определенных областях исследований и менее значительно влияют на результат в других. Для того, чтобы рассмотреть разные случаи были выбраны четыре метода из разных подходов: два относятся к текстурной обработке изображений, два – к снижению уровня шума и улучшению качества изображений.

Сегментация изображений. Для сегментации изображений позвоночника применяются несколько разных методов, их можно разделить на три группы: классические методы компьютерного зрения, методы машинного обучения и нейронные сети.

В статье [11] предложен полуавтоматический алгоритм сегментации позвонков с использованием классических методов компьютерного зрения. Авторами был подробно описан алгоритм работы, проведена всесторонняя оценка результатов, а также сравнение с другими работами. Исходные данные содержали 63 изображения, всего 419 позвонков. К преимуществам данного метода можно отнести высокую скорость обработки изображений, а также возможность применения на изображениях с различными патологиями, такими как: переломы, сколиоз, метастазы и т.д. Среди недостатков – необходимость участия человека в процессе обработки изображений и меньшая точность сегментации, чем в других исследованиях: $88.2 \pm 1.9\%$ на T2-взвешенных изображениях. Для оценки результата использовался коэффициент Дайса.

Для задачи сегментации позвонков в работе [12] предлагается использовать метод опорных векторов, на вход которому подаются такие параметры области позвонка как интенсивность вокселя, среднее значение яркости пикселя, дисперсия, гауссиан, медиана, оператор Собеля, и др. Авторы проверили работу своего метода на различных наборах данных, содержащих изображения, полученных с помощью КТ и МРТ (как T1, так и T2-взвешенные изображения). Наилучший результат был достигнут для КТ изображений – 98%. Применение алгоритма на T2-взвешенных МРТ

изображениях без мышечной ткани позволило достичь значения 96.7%, с мышечной тканью – 91.7%. Для оценки результата использовался коэффициент Дайса.

Метод, предложенный в статье [13] основан на объединении каскада классификаторов и активной модели внешнего вида на основе метода главных компонент. Авторы провели масштабное сравнение результатов с результатами других исследований, а также с результатами ручной обработки пяти экспертов. Использовался набор данных из 50 МРТ последовательностей, который был искусственно увеличен с помощью метода тонких пластинчатых сплайнов (Thin Plate Splines, TPS). В данном исследовании была получена точность (accuracy) 91.37%.

В 2017 году, была разработана многозадачная сверточная нейронная сеть, позволяющая не только сегментировать позвонки, но также классифицировать каждый позвонок по уровню его деформации, определять возможное заболевание и рассчитывать объем межпозвоночного диска [14]. Кроме этого, сеть может выделять «горячие точки» – места, которые требуют особого внимания специалистов, например, места деформаций позвонка или межпозвоночного диска. Данная нейронная сеть была основана на нескольких уже существующих методах, однако они были существенно улучшены.

Следующий метод, предложенный в статье [15], использует нейронную сеть с глубоким обучением на основе архитектуры U-Net для мультиклассовой сегментации позвоночника. После этого для каждого изображения строилась кривая, описывающая изгиб позвоночника, находились центры межпозвоночных дисков. Затем выделенные области межпозвоночных дисков в сагиттальной и аксиальной проекциях подавались на вход классификатору, которые предсказывал наличие стеноза и его тяжесть. В данном исследовании подробно описаны архитектуры всех используемых нейронных сетей, также авторы утверждают, что данный метод надежен при обработке практически всех изображений за исключением случаев тяжелого сколиоза, сросшихся

позвонков или артефактов на изображении (движения или металлических). Полученный коэффициент Дайса для сегментации в среднем равен 93%.

В работе [16] предложен итеративный алгоритм сегментации, основанный на полностью сверточной многозадачной нейронной сети, которую можно разделить на три компонента: сеть для сегментации на основе архитектуры U-Net, дополненная специальным блоком памяти экземпляра, идентификационная подсеть, для классификации позвонков и подсеть классификации полноты, позволяющая различать полностью и частично видимые позвонки. Итерационный подход заключается в том, что изображение сегментируется в определенной последовательности: снизу-вверх или сверху-вниз. Используется специальное скользящее окно, которое проходит по изображению в выбранном направлении с заданным шагом. Блок памяти экземпляра позволяет запомнить какие позвонки были сегментированы на предыдущих итерациях, что позволяет нейронной сети игнорировать их и сфокусироваться на следующем, еще не сегментированном позвонке. Для сегментации использовались КТ и МРТ изображения. Авторами было проведено сравнение результатов с результатами других исследований, а также сравнение работы алгоритма на изображениях исследований различных отделов позвоночника. Наибольший коэффициент Дайса был получен для КТ изображений поясничного отдела – $96.5 \pm 0.8\%$, для МРТ изображений поясничного отдела было получено – $94.4 \pm 3.3\%$

Анализ существующих исследований показал, что наибольшую точность обеспечивает применение алгоритмов с использованием сверточных нейронных сетей. Также, многие авторы берут за основу своих разработок архитектуру U-Net.

В данной работе рассматривается задача сегментации МРТ изображений позвоночника. Выделяется два основных этапа: предварительная обработка изображений и сегментация с использованием сверточных нейронных сетей.

Исходные данные

Для сегментации использовались МРТ изображения поясничного отдела позвоночника в сагиттальной проекции, полученные из трех источников:

SpineSagT2Wdataset3. Открытый набор данных, собранный для конкурса, проводимого в Китае в 2019 году для студентов. Всего в нем содержится 210 файлов формата NiFTY (195 обучающая выборка и 15 тестовая). Каждый файл представляет собой последовательность из 11-14 изображений поясничного отдела позвоночника в сагиттальной проекции, полученных с помощью магнитно-резонансной томографии. Какой-либо информации о пациентах не предоставлено.

Области, не представляющие интереса на изображениях затемнены. Также указано, что маски для обучающей выборки были созданы вручную специалистами.

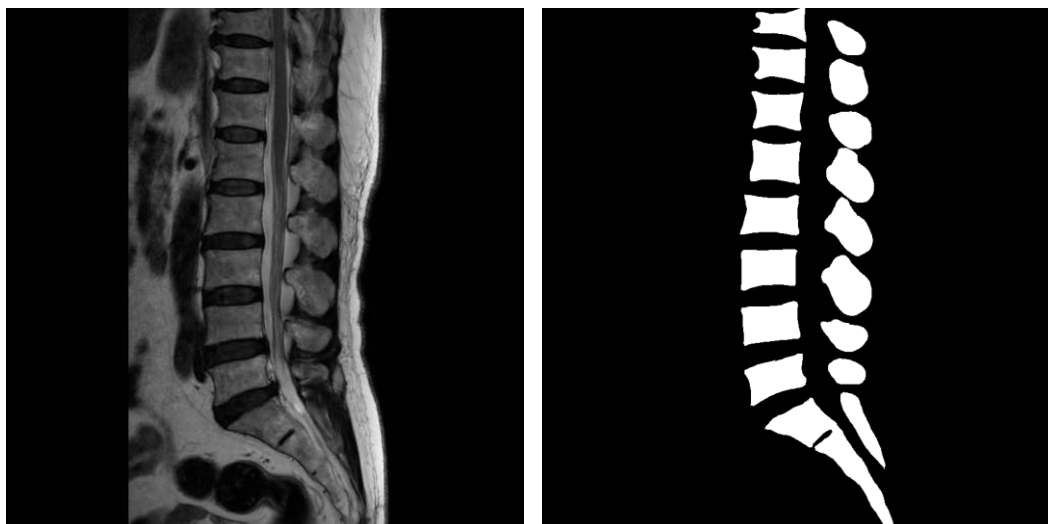


Рисунок 1. Пример изображения и маски из набора данных SpineSagT2Wdataset3

Dataset 11: High anisotropy MRIs of the lower back. Открытый набор данных, который использовался для исследования [17] немецких разработчиков.

Набор содержит 17 анонимизированных изображений формата DICOM в различных проекциях. В нем присутствуют как изображения всего позвоночника, так и исключительно поясничного отдела.

Изображения были получены из различных больниц, с использованием разного оборудования, поэтому они довольно сильно различаются (рис. 2).

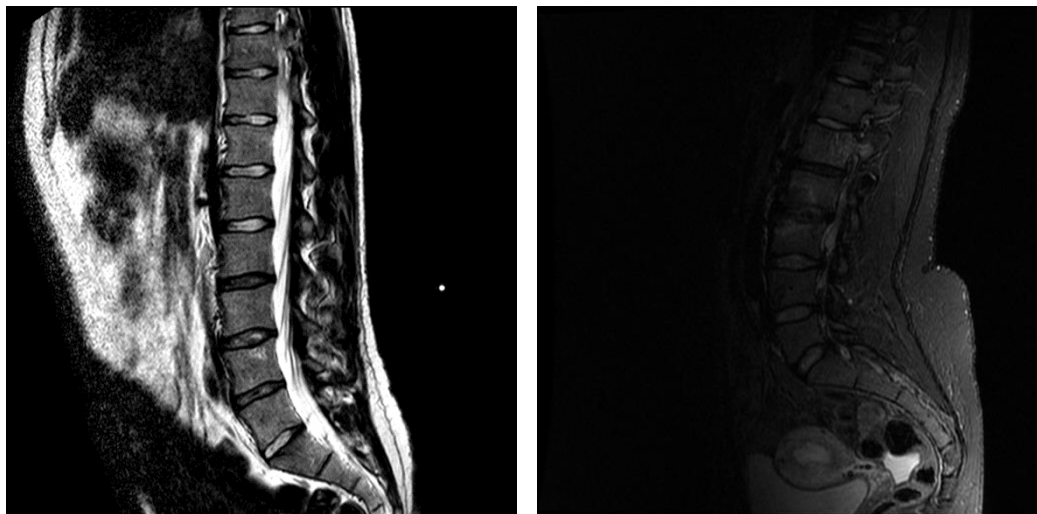


Рисунок 2. Примеры изображений из набора данных Dataset 11

В данном наборе содержится информация о пациентах, а именно: возраст (от 21 до 74, в среднем 40), пол (11 женщин и 6 мужчин), состояние позвоночника (здоров – 2, сколиоз – 1, перелом позвонка – 3, спондилолистез – 3, 9 – другие патологии, не диагностируемые по сегментации позвоночника) и др.

ФГБУ «НМИЦ им. В. А. Алмазова». Для исследования было выделено 130 изображений в различных проекциях. Данные анонимизированы и информации о пациентах нет.

Глава 1. Применение методов обработки изображений

Для повышения качества сегментации часто используются различные способы обработки изображений. К ним относятся как простые операции изменения размера или цветовой палитры, так и применение более сложных фильтров.

В данной работе предлагается использовать несколько методов текстурного анализа: статистический – матрица смежности уровней серого Харалика и спектральный – фильтр Габора. Кроме того, были использованы нелинейный геометрический фильтр Кримменса и простой фильтр нерезкое маскирование, который часто применяется при обработке изображений. Основной задачей последних двух методов является повышение качества изображения.

1.1. Матрица смежности уровней серого

Построение матрицы смежности уровней серого (Gray Level Co-occurrence Matrix – GLCM) – один из самых ранних методов извлечения текстурных признаков, который был предложен Хараликом, Шанмуганом и Динстейном в 1971 году [18, 19]. Основой данного метода является предположение, что информация о текстуре изображения заключается в пространственном отношении оттенков серого, содержащихся в изображении, друг к другу.

Каждый элемент матрицы GLCM вычисляется по формуле:

$$G_{(\Delta x, \Delta y)}(i, j) = \sum_{y=0}^{h-\Delta y} \sum_{x=0}^{w-\Delta x} F,$$

$$F = \begin{cases} 1, & \text{если } I(x, y) = i \text{ и } I(x + \Delta x, y + \Delta y) = j, \\ 0, & \text{иначе.} \end{cases}$$

где i, j – яркость пикселя, $(\Delta x, \Delta y)$ – заданная окрестность, $I(x, y)$ – яркость пикселя с координатами (x, y) , w, h – ширина и высота исходного изображения.

В зависимости от значений Δx и Δy можно получить различные матрицы смежности, выбирая для сравнения яркости «соседа» под различными углами (рис. 3).

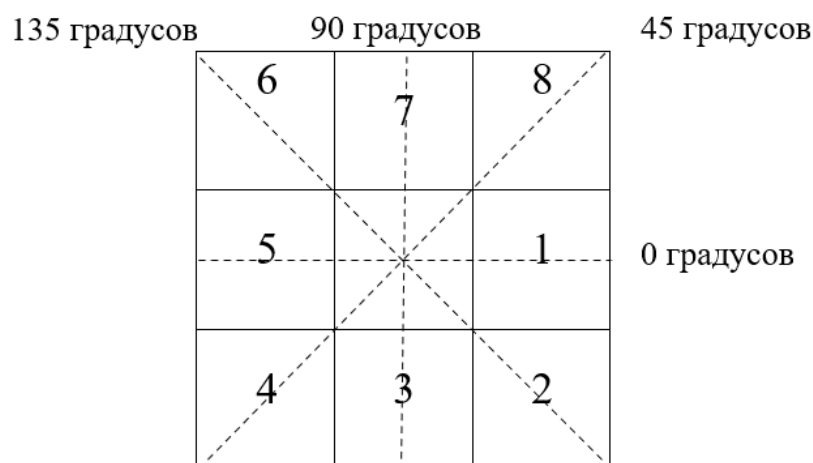


Рисунок 3. Соседние пиксели

Обычно используются матрицы $G_{(1,0)}$, $G_{(1,1)}$, $G_{(0,1)}$ и $G_{(-1,-1)}$.

Полученные матрицы нормируются:

$$P_{(\Delta x, \Delta y)} = W G_{(\Delta x, \Delta y)}(i, j),$$

$$W = \frac{1}{(w - \Delta x)(h - \Delta y)}.$$

В результате получается четыре матрицы. Можно сказать, что полученные матрицы представляют собой оценки совместных плотностей распределения вероятностей того, что два пикселя с яркостями i и j окажутся соседями. Полученные матрицы используются для вычисления 14 характеристик Харалика. В данной работе используются лишь некоторые, приведенные ниже:

Контраст:

$$Contrast = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} (i - j)^2 P_{i,j}$$

Несходство:

$$Dissimilarity = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} |i - j| P_{i,j}$$

Однородность:

$$Homogeneity = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2}$$

Энтропия:

$$Entropy = - \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} P_{i,j} \ln P_{i,j}$$

Второй угловой момент:

$$ASM = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} P_{i,j}^2$$

Энергия:

$$Energy = \sqrt{ASM}$$

Где N – размерность матрицы $P_{(\Delta x, \Delta y)}$

1.2. Геометрический фильтр Кримминса

Геометрический фильтр, разработанный Т.Р. Кримминсом основан на применении к изображению итеративного алгоритма. Он был разработан для удаления с радиолокационных изображений шума типа «соль-перец» [20].

Можно сказать, что это одномерный алгоритм, который применяется последовательно в четырех разных направлениях на двумерном изображении: горизонтальном, вертикальном и двух диагональных направлениях. Значение интенсивности в центре каждого окна размером 3×3 сравнивается с ближайшими окружающими его пикселями. В зависимости от результатов сравнения, значение интенсивности центрального пикселя увеличивается, либо уменьшается.

Рассмотрим алгоритм фильтра для вертикального направления. Пусть на рисунке 4 представлена матрица изображения.

a	e	i	m	q
b	f	j	n	r
c	g	k	o	s
d	h	l	p	t

Рисунок 4. Пример применения геометрического фильтра. Матрица изображения

Алгоритм для пикселя b имеет вид:

Для каждой итерации:

1. Настройка темных пикселей:
 - a. if $a \geq b + 2$ then $b = b + 1$
 - b. if $a > b$ and $b \leq c$ then $b = b + 1$
 - c. if $c > b$ and $b \leq a$ then $b = b + 1$
 - d. if $c \geq b + 2$ then $b = b + 1$
2. Настройка светлых пикселей:
 - a. if $a \leq b - 2$ then $b = b - 1$
 - b. if $a < b$ and $b \geq c$ then $b = b - 1$
 - c. if $c < b$ and $b \geq a$ then $b = b - 1$
 - d. if $c \leq b - 2$ then $b = b - 1$

В результате разница между значениями интенсивностей соседних пикселей уменьшается, а изображение – сглаживается. Уровень сглаживания зависит от количества итераций.

1.3. Фильтр «Нерезкое маскирование»

Нерезкое маскирование (Unsharp masking) – метод, который используется для повышения резкости изображения [21]. Основная идея заключается в использовании размытого («нерезкого») изображения для создания маски, которая впоследствии объединяется с исходным изображением. В результате получается более резкое изображение, чем оригинал.

Для построения отфильтрованного изображения используется следующая формула:

$$f_{sharp}(x, y) = f(x, y) + kg(x, y)$$

$$g(x, y) = f(x, y) - f_{smooth}(x, y)$$

где $f(x, y)$ – исходное изображение, $f_{smooth}(x, y)$ – исходное изображение с применением сглаживания (по Гауссу), $g(x, y)$ – изображение края, k – константа.

1.4. Фильтры Габора

Фильтр Габора [5, 22] представляет собой линейный фильтр, который используется в задачах обнаружения краев, анализа текстур, извлечения объектов и т. д. Было показано, что эти фильтры обладают оптимальными свойствами локализации как в пространственной, так и в частотной области и поэтому хорошо подходят для задач сегментации текстур. Данный фильтр относится к специальному классу полосовых фильтров, следовательно, допускает определенную «полосу» частот и отклоняет другие.

Зачастую используется не один фильтр Габора, а набор, построенный с использованием различных ориентаций.

В двумерном случае, данный фильтр можно рассматривать как синусоидальный сигнал с определенными значениями частоты и ориентации, модулированный гауссовой волной. Фильтр можно использовать как в виде комплексного числа, так и разделив его на две отдельные компоненты: действительную и мнимую.

Комплексное:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right).$$

Действительная часть:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \psi\right).$$

Мнимая часть:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'}{\lambda} + \psi\right).$$

Где

$$x' = x \cos \theta + y \sin \theta$$

$$y' = -x \sin \theta + y \cos \theta$$

В приведенных выше уравнениях,

λ – длина волны синусоидальной составляющей;

θ – ориентация нормали параллельных полос функции;

ψ – сдвиг фаз;

σ – стандартное отклонение гауссовой огибающей;

γ – коэффициент сжатия, отвечающий за эллиптичность функции.

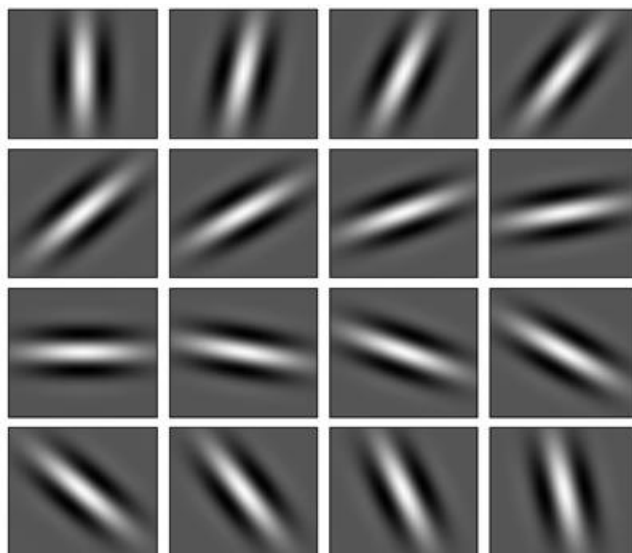


Рисунок 5. Пример фильтров Габора

Глава 2. Архитектуры нейронных сетей для сегментации изображений

Наиболее перспективным направлением нейронных сетей в медицине считается применение сверточных нейронных сетей для обработки изображений. Так как для обучения моделей требуются очень большие наборы данных, которые трудно собрать в этой области, некоторое время применение искусственного интеллекта для обработки медицинских изображений было затруднительным.

В 2015 году эта проблема была решена созданием архитектуры U-Net, которая обеспечивала высокую точность сегментации при меньших наборах данных. В последствие на основе U-Net было разработано множество различных архитектур, например: ResUnet, ResUnet++, Nested U-Net, UNet 3+ и т.д.

2.1. U-Net

U-Net состоит из двух частей, сужающейся (энкодер) и расширяющейся (декодер) (рис. 6).

Энкодер представляет собой типичную архитектуру сверточной сети и состоит из основных блоков, включающих две операции свертки, за каждой из которых следует операция активации. За каждым блоком следует операция подвыборки по максимальному значению. Строение блока представлено на рисунке 8 (а).

Декодер состоит из шагов обратной свертки (деконволюции), которая уменьшает количество каналов после чего следует конкатенация с обрезанной картой признаков от соответствующей части энкодера и основной блок. При каждой операции свертки происходит потеря пограничных пикселей, поэтому выполняется обрезка.

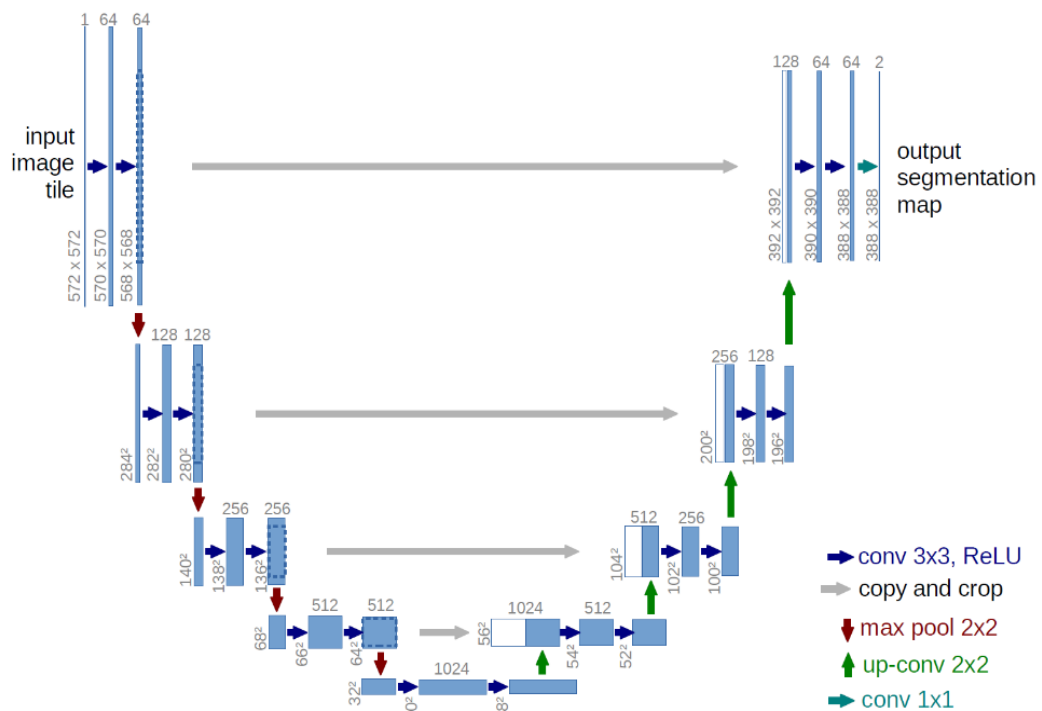


Рисунок 6. Архитектура U-Net (пример для 32x32 пикселей в самом низком разрешении) [23]

2.2. Residual U-Net

ResUnet – модель для семантической сегментации, вдохновленная глубоким остаточным обучением (ResNet) и U-Net, объединяющая преимущества двух этих архитектур.

Основной особенностью ResNet является использование «остаточной» структуры обучения. Для этого использовались соединения быстрого доступа (shortcut connections), которые пропускают некоторое количество слоев и выполняют сопоставление идентификаторов. Введение подобных методов позволило избежать увеличения ошибок обучения при увеличении количества слоев (так называемая проблема деградации), а значит позволило строить более глубокие сети с более высокой точностью. Кроме этого, это позволяет проектировать сети с меньшим количеством параметров без потери точности.

Архитектура ResUnet представлена на рисунке 7. Она состоит из трех частей: энкодера, моста (bridge) и декодера. На рисунке 8 представлены основные блоки U-Net и ResUnet.

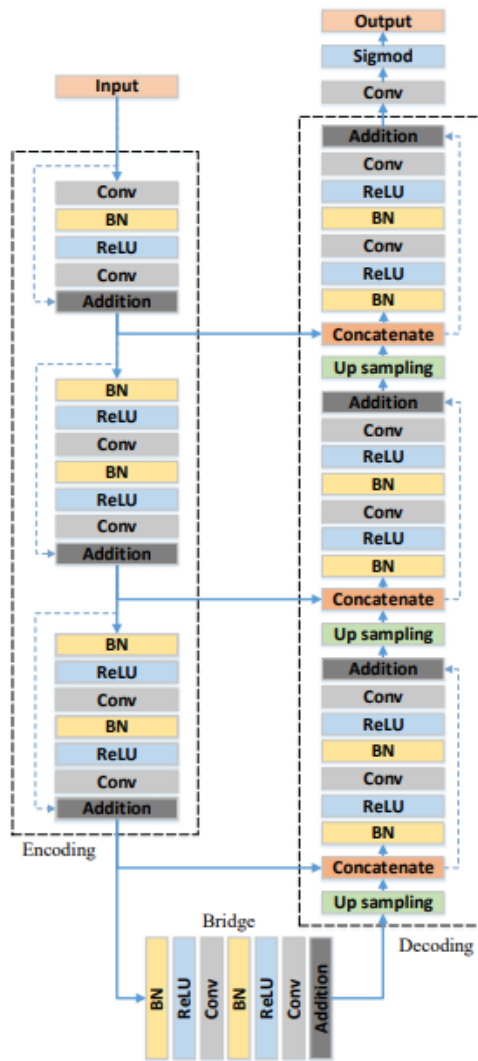


Рисунок 7. Архитектура ResUnet [24]

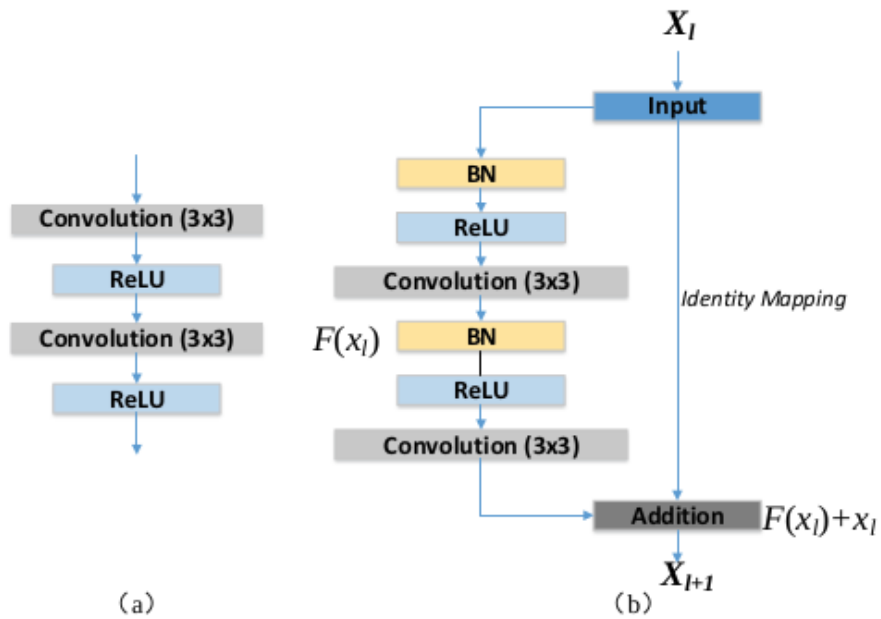


Рисунок 8. Блоки нейронных сетей. (а) Простая нейронная единица, используемая в U-Net, и (б) остаточная единица с сопоставлением идентичности, используемая в ResUnet [24].

2.3. ResUnet++

В 2019 году была представлена модель, более эффективная и производительная чем U-Net и ResUnet.

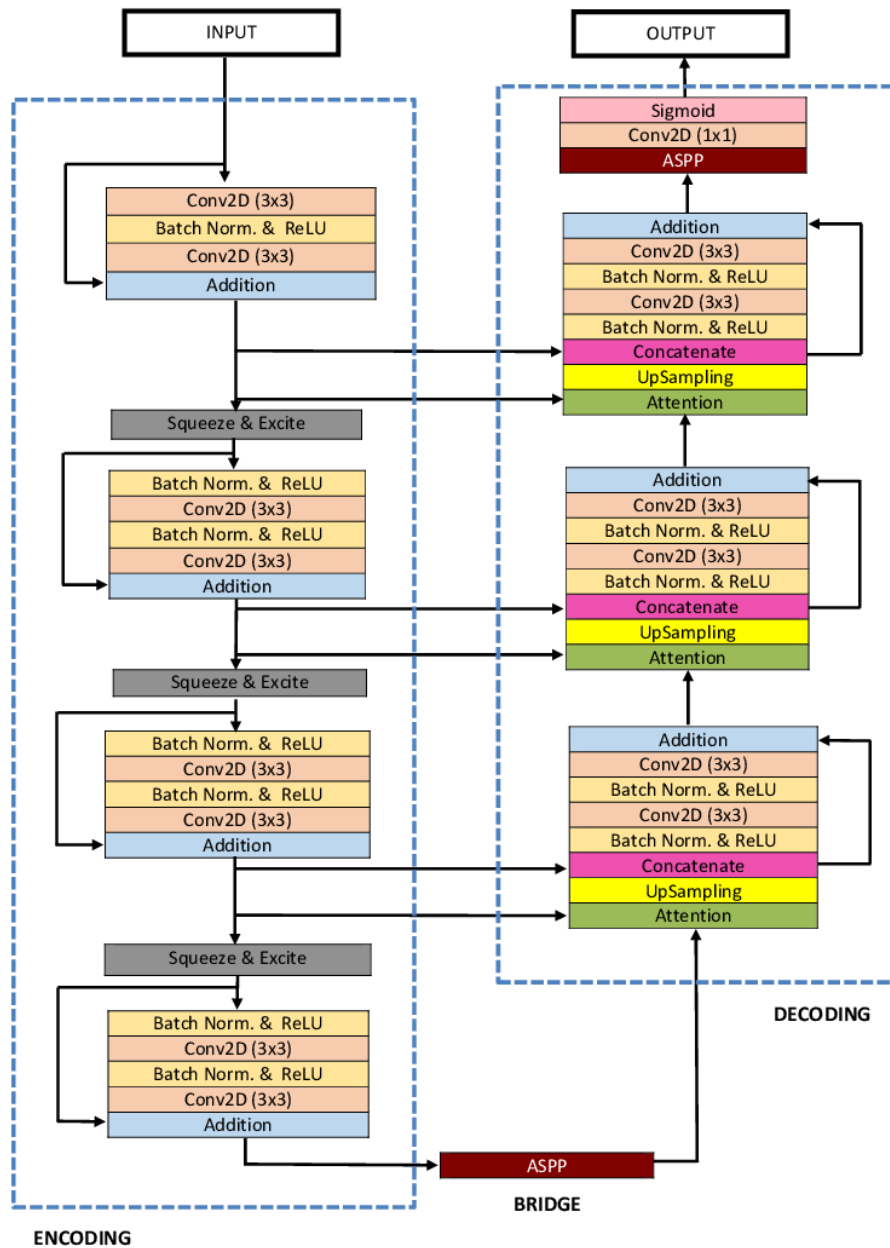


Рисунок 9. Архитектура ResUnet++ [25]

ResUnet++ (рис. 9) основана на архитектуре ResUNet, однако, она была дополнена новыми блоками:

1. блок сжатия и стимуляции (squeeze and excitation block) – позволяет распределить вычислительные ресурсы в пользу более информативных частей сигнала за счет присвоения каналам весов;

2. блок ASPP (Atrous Spatial Pyramidal Pooling) – позволяет эффективно захватывать многомасштабную информацию за счет параллельного применения сверток с различными скоростями расширения для увеличения поля зрения;

3. блок внимания (attention block) – используется для выделения фрагментов данных, которые требуют более детальной обработки нейронной сети.

2.4. UNet++

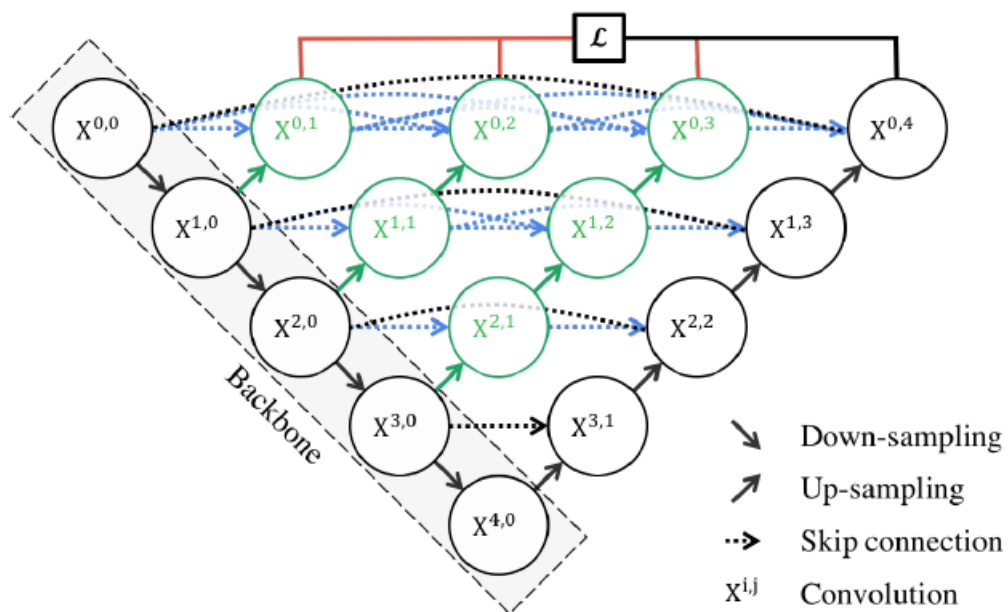


Рисунок 10. Архитектура U-Net++ [26]

Особенностью UNet++ является использование идеи плотных (компактно соединенных) блоков, используемых в DenseNet. Такие блоки позволяют получать данные от всех предыдущих слоев и передавать свои карты признаков последующим слоям.

Этот подход позволяет сократить семантический разрыв между картами признаков кодера и декодера и, как следствие, повысить точность сегментации.

Кроме того, в данной архитектуре реализовано «глубокое наблюдение» (Deep supervision), которое позволяет настраивать сложность модели отсекая слои.

Глава 3. Разработка и реализация алгоритма автоматической сегментации

Алгоритм состоит из трех основных этапов: начальной обработки изображений, дополнительной обработки, обработка нейронной сетью.

Начальная обработка изображений. Данный этап включает в несколько пунктов. Так как изображения получены из разных источников, они имеют различные форматы, размеры и качество. Первоначально изображения приводятся к одному формату и размеру. Затем, если это необходимо, корректируется яркость, затемняются области, не представляющие интереса, создается маска.

Дополнительная обработка изображений. Здесь к изображениям применяется один из выбранных для исследования методов обработки (основанный на матрице смежности уровней серого Харалика, фильтры Габора, геометрический алгоритм Кримминса или «нерезкое маскирование»). Так же есть возможность оставить изображения без изменения.

Обработка нейронной сетью. Если обучения модели еще не проводилось, то изображения разделяются на обучающую, валидационную и тестовую выборки, после чего подаются в нейронную сеть для обучения. После обучения оцениваются результаты. В случае, когда обучение уже проведено модель предсказывает для полученных изображений маски.

Схематичное описание алгоритма представлено на рисунке 11.



Рисунок 11. Схема алгоритма

Алгоритм был реализован на языке программирования Python версии 3.6 с использованием библиотек:

1. для предобработки изображений – Pydicom, OpenCV, Pandas, Matplotlib, Scikit-image, NumPy, Pillow;
2. для создания и обучения моделей – Scikit-learn, Keras и TensorFlow.

Разработка проводилась на стационарном компьютере со следующими характеристиками:

- ЦПУ: Intel(R) Core (TM) i5-7600K CPU @ 3.80GHz (4 CPUs);
- ОЗУ: 16 ГБ;
- ГПУ: NVIDIA GeForce GTX 1070 Ti, с объемом памяти 8 Гб.

Для обучения моделей ResUnet++ и Unet++ использовался сервис Google Colaboratory.

3.1. Предварительная обработка изображений

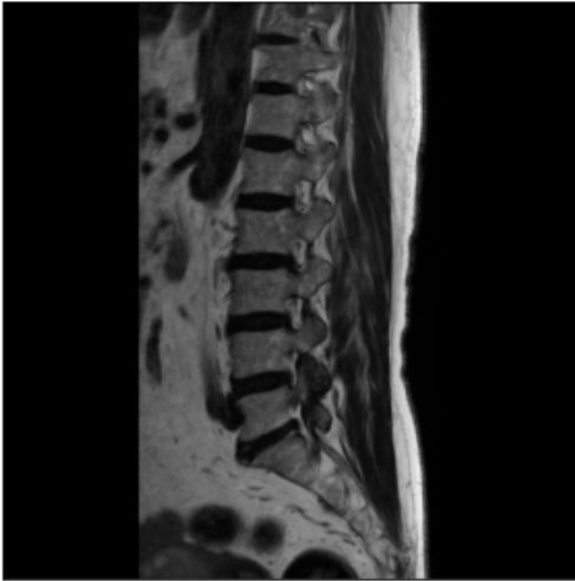
Изначально все изображения были переведены из форматов NiFTY (SpineSagT2Wdataset3) и DICOM (Dataset 11 и изображения, предоставленные ФГБУ «НМИЦ им. В. А. Алмазова») в png. Для изображений из второго и третьего источников были затемнены края изображений и созданы вручную маски. Также, для некоторых изображений из источника 2 потребовалась коррекция яркости.

После завершения начальной обработки к каждому изображению был применен набор фильтров. Рассмотрим результаты обработки:

Геометрический фильтр Кримминса

Чтобы не потерять важные детали, была проведена одна итерация геометрического алгоритма.

Исходное изображение



Изображение с фильтрацией

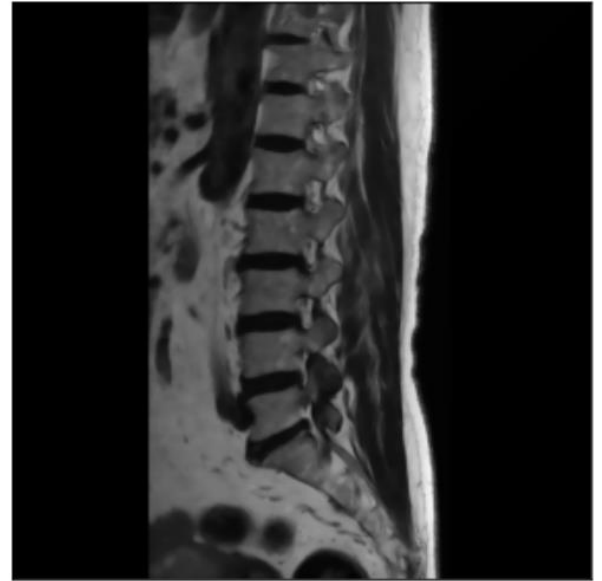
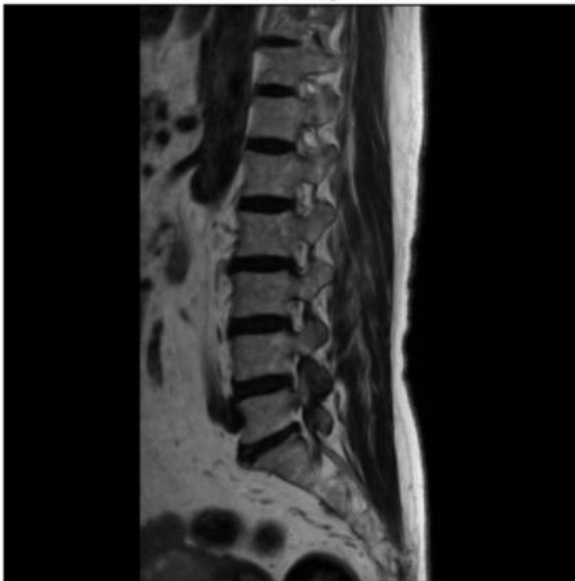


Рисунок 12. Результат обработки изображения с использованием геометрического фильтра

Фильтр «Нерезкое маскирование»

Исходное изображение



Изображение с фильтрацией

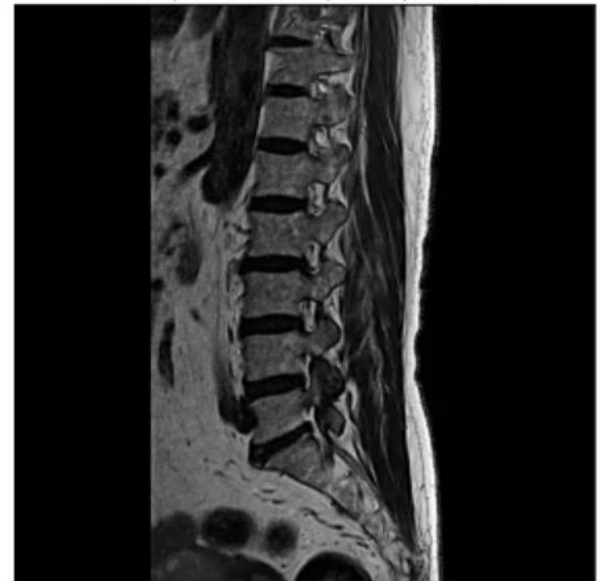


Рисунок 13. Результат обработки изображения с использованием фильтра «Нерезкая маска»

Матрица смежности уровней серого

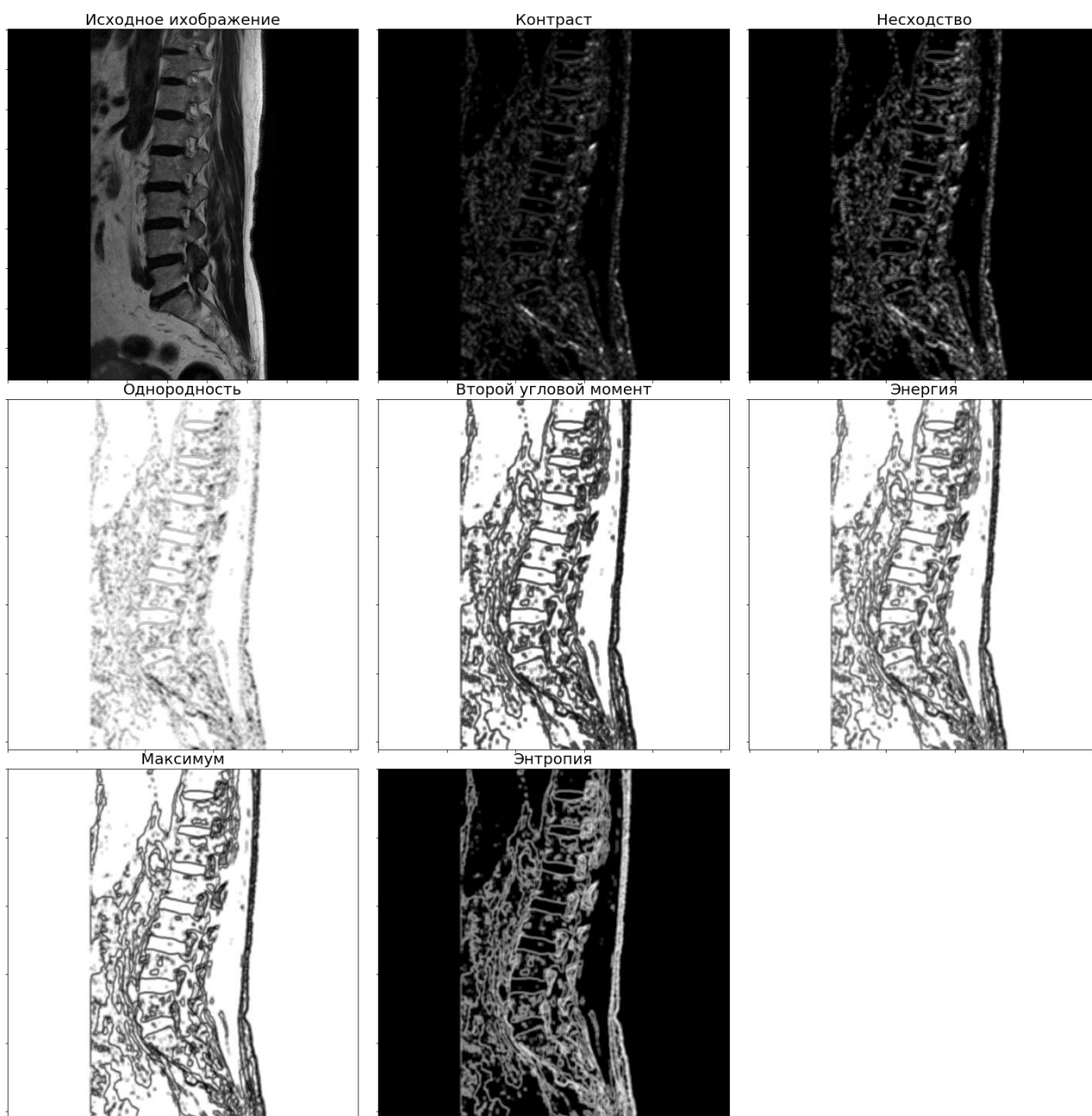


Рисунок 14. Результат обработки изображения с использованием параметров Харалика

Вейвлет преобразование Габора

Было построено несколько фильтров с различными значениями параметров (табл. 1).

Таблица 1. Параметры для фильтров Габора

	Размер ядра	λ	ψ	σ	θ
№1	(16, 16)	6.25	0.5	3	$[0, \pi]$
№2	(16, 16)	6.25	0.1	3	$[0, \pi]$
№3	(8, 8)	6.25	0	3.5	$[0, \pi]$

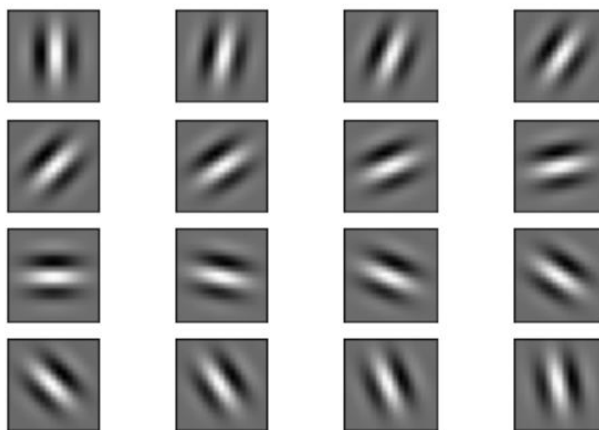


Рисунок 15. Набор фильтров №1

На рисунке 16 представлены полученные изображения для каждого фильтра.

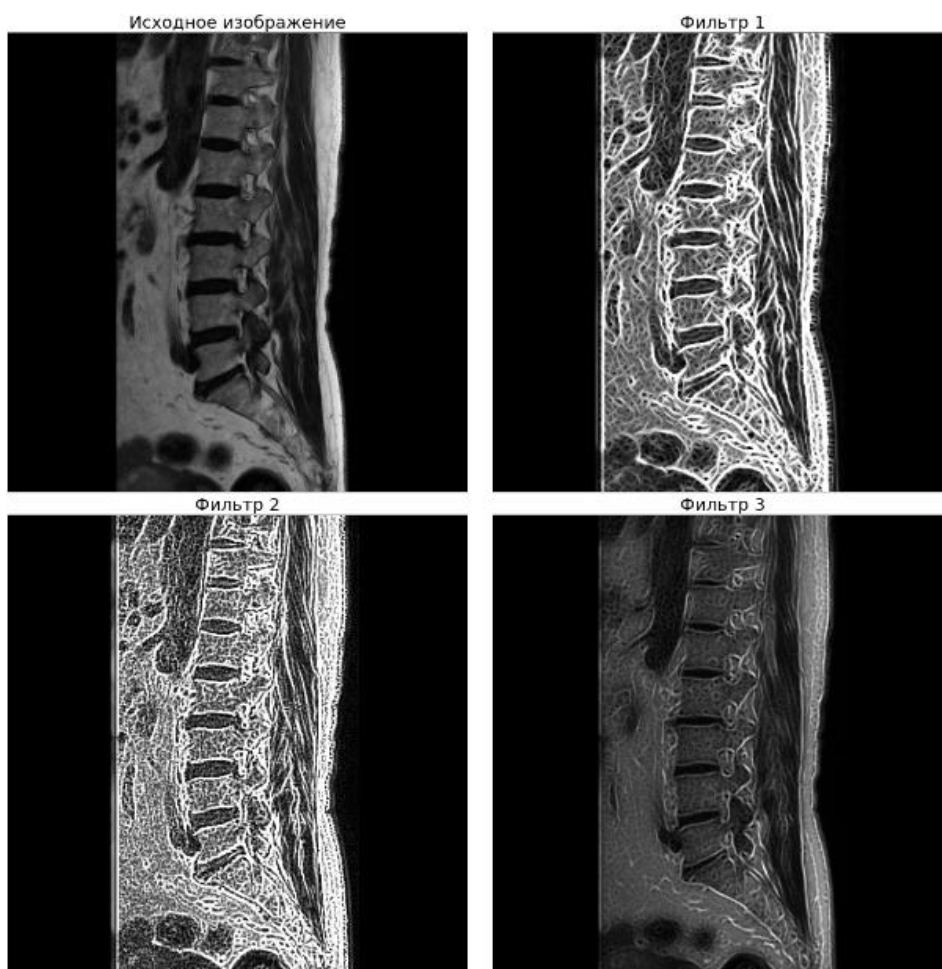


Рисунок 16. Результат обработки изображения

3.2. Сегментация изображений

В качестве функции активации в моделях используется ReLU (rectified linear unit), которая реализует простой пороговый переход в нуле.

В качестве метрик для обучения сети и оценки качества сегментации, был выбран коэффициент Дайса (Dice coefficient) (коэффициент Сёренсена — Sorensen–Dice coefficient), а также функция потерь.

Коэффициент Дайса

Коэффициент Dice определяется следующей формулой:

$$Dice(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|}$$

где X является предсказанием, а Y - правильно размеченной маской на текущем объекте. $|X|$ означает мощность множества X (количество элементов в этом множестве), а \cap – пересечение X и Y .

Функция потерь

В качестве функции потерь была выбрана:

$$loss = 1 - Dice(X, Y).$$

Кроме того, для каждой модели были посчитаны значения precision, recall и F1- score.

Precision (точность) – показывает, какая часть пикселей, отнесенная к классу маски, действительно принадлежит ей.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

Recall (чувствительность) – показывает, какую часть пикселей, принадлежащих маске позвонка, обнаружил алгоритм.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

F1- score гармоническое среднее точности и чувствительности.

$$F1 = 2 \frac{Precision * Recall}{Precision + Recall}$$

Были построены и обучены четыре модели: U-Net, ResUnet, ResUnet++ и UNet++. Все модели, кроме UNet++, были модифицированы путем добавления дополнительных блоков, для увеличения глубины, начальный вектор признаков моделей содержит 16 компонент. UNet++ была реализована в классическом виде.

Объем обучающей выборки составил 2130 изображений, валидационной – 480, тестовой – 193.

3.2.1. Обучение без дополнительной предварительной обработки.

U-Net

Обучение длилось 80 эпох и заняло 2 часа 57 минут 16 секунд.

На рисунке 17 представлены графики потерь и коэффициента Dice на обучающей и тестовой выборках. На рисунке 18 – графики значений точности и чувствительности.

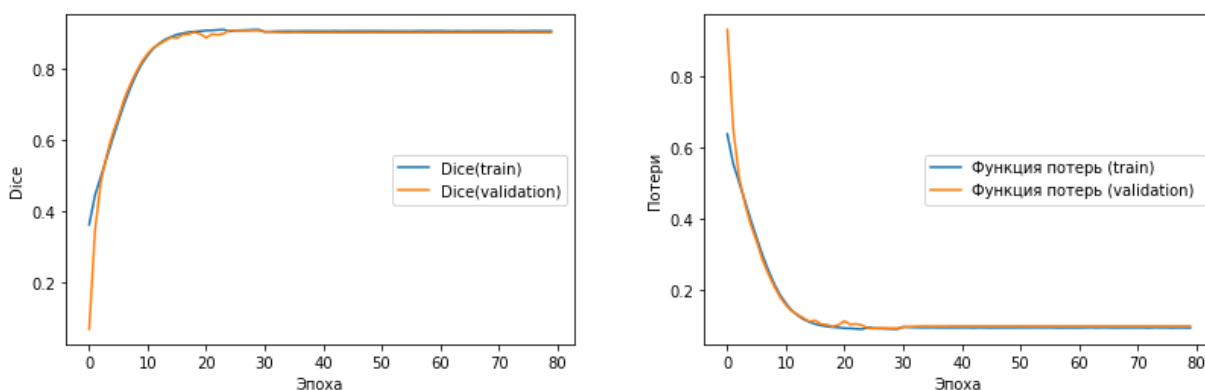


Рисунок 17. Значения коэффициента Дайса и функции потерь на обучающей и валидационной выборке

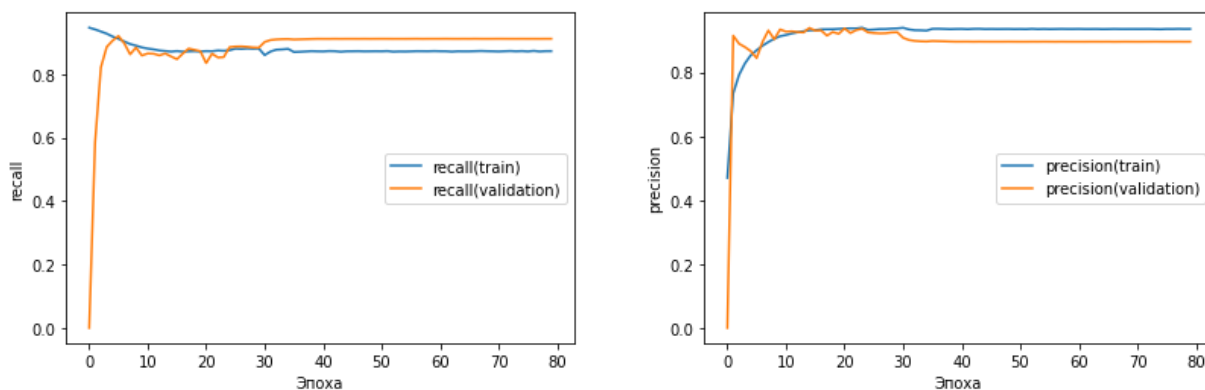


Рисунок 18. Значения чувствительности и точности на валидационной и тестовой выборках

Коэффициент Дайса равен 0.9091 для обучающей и 0.9080 для тестовой выборки. Минимальное значение коэффициента Дайса на тестовой выборке равно 0.6578, максимальное – 0.9586.

ResUnet

Обучение длилось 80 эпох и заняло 2 часа 52 минуты 54 секунд.

На рисунке 19 представлены графики потерь и коэффициента Dice на обучающей и тестовой выборках. На рисунке 20 – графики значений точности и чувствительности.

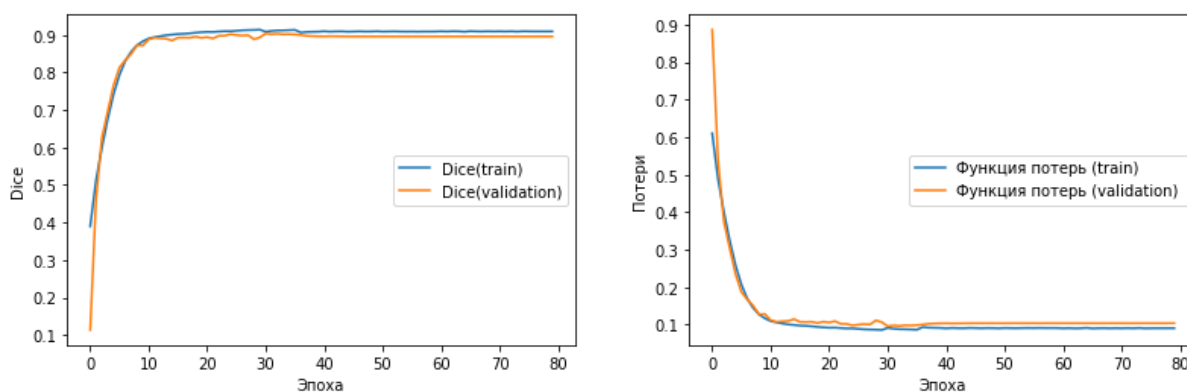


Рисунок 19. Значения коэффициента Дайса и функции потерь на обучающей и валидационной выборке

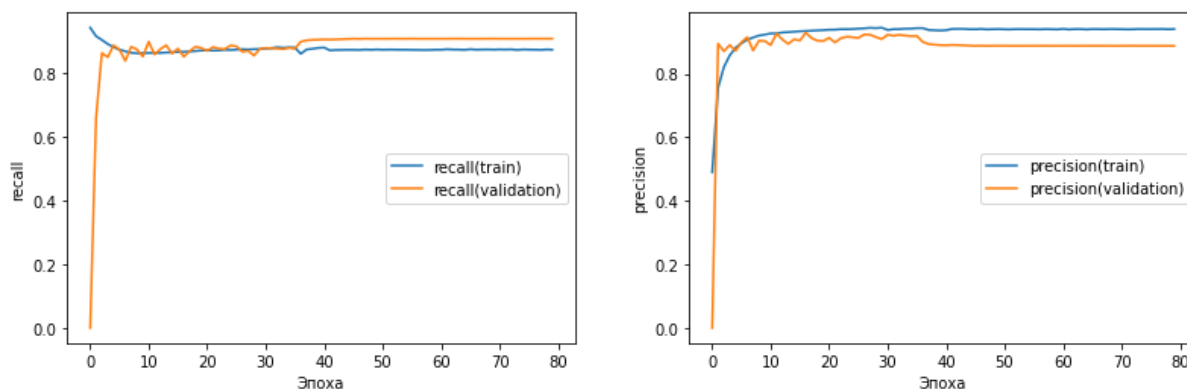


Рисунок 20. Значения чувствительности и точности на валидационной и тестовой выборках

Коэффициент Дайса равен 0.9119 для обучающей и 0.9017 для тестовой выборки. Минимальное значение коэффициента Дайса на тестовой выборке равно 0.6467, максимальное – 0.9622.

ResUnet++

Обучение длилось 39 эпох и заняло 5 часов 33 минуты 42 секунд.

На рисунках 21-22 представлены графики потерь и коэффициента Dice на обучающей и тестовой выборках, а также графики значений точности и чувствительности.

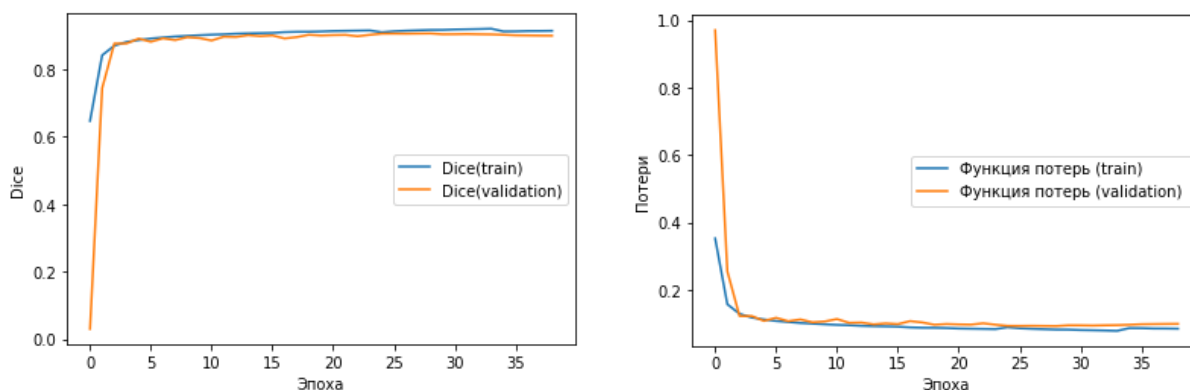


Рисунок 21. Значения коэффициента Дайса и функции потерь на обучающей и валидационной выборке

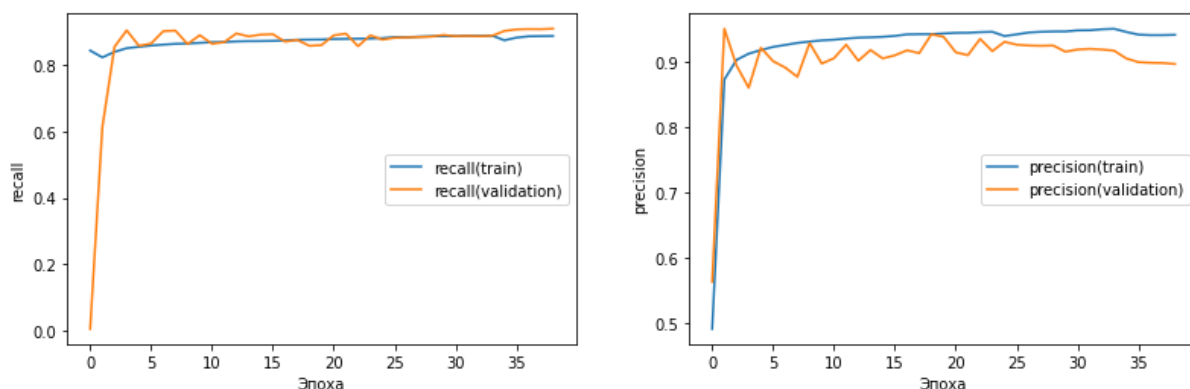


Рисунок 22. Значения чувствительности и точности на валидационной и тестовой выборках

Значение коэффициента Дайса равно 0.9198 для обучающей и 0.9108 для тестовой выборки. Минимальное значение коэффициента Дайса на тестовой выборке равно 0.7860, максимальное – 0.9645.

UNet++

Обучение длилось 75 эпох и заняло 13 часов 11 минут 57 секунд.

На рисунке 23 представлены графики потерь и коэффициента Dice на обучающей и тестовой выборках. На рисунке 24 – графики значений точности и чувствительности.

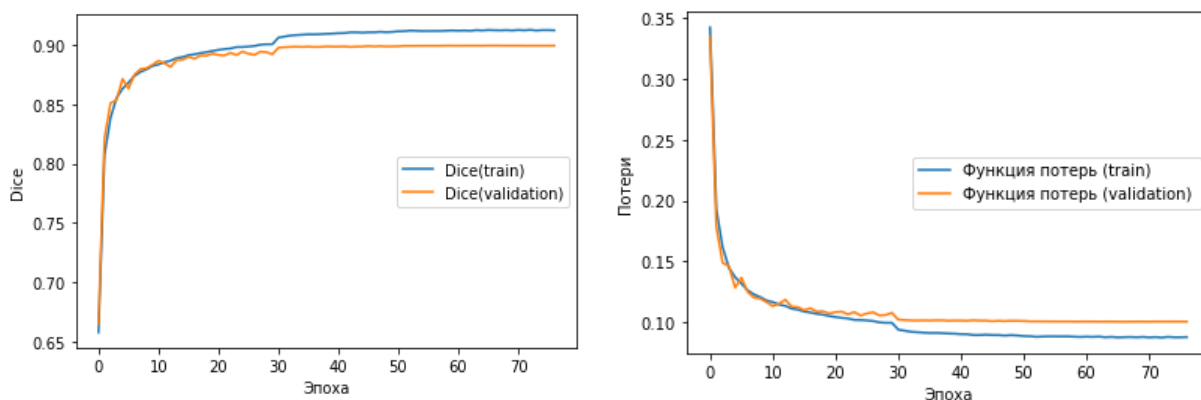


Рисунок 23. Значения коэффициента Дайса и функции потерь на обучающей и валидационной выборке

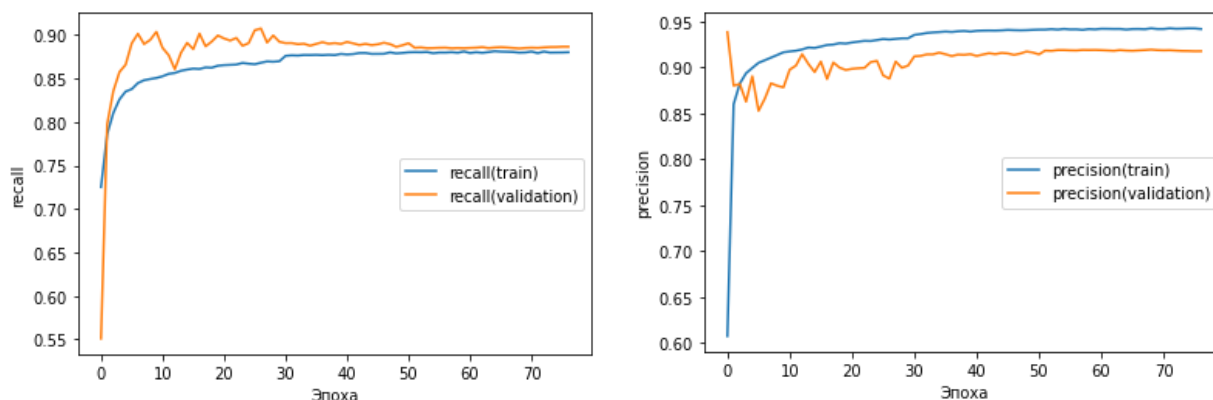


Рисунок 24. Значения чувствительности и точности на валидационной и тестовой выборках

Значение коэффициента Дайса равно 0.9168 для обучающей и 0.9044 для тестовой выборки. Минимальное значение коэффициента Дайса на тестовой выборке равно 0.6280, максимальное – 0.9656.

В таблице 2 представлены основные результаты оценки точности сегментации на тестовой выборке.

Таблица 2. Оценка точности сегментации на тестовой выборке без предварительной обработки

	<i>Dice</i>	<i>Dice</i> _{мин.}	<i>Dice</i> _{макс.}	<i>Precision</i>	<i>Recall</i>	<i>F1</i>
U-Net	0.9080	0.6578	0.9586	0.9329	0.8832	0.9074
ResUnet	0.9017	0.6467	0.9622	0.9314	0.8735	0.9015
ResUnet++	0.9108	0.7860	0.9645	0.9346	0.8849	0.9091
UNet++	0.9044	0.6280	0.9656	0.9288	0.8798	0.9036

3.2.2. Обучение с применением дополнительной предварительной обработки

В таблице 3 представлены результаты обучения моделей при различной предварительной обработке.

Таблица 3. Результаты обучения с использованием дополнительной предварительной обработки

Фильтр		Значение коэффициента Дайса	
		U-Net	ResUnet
Без фильтра		0.9080	0.9018
Признаки Харалика	Энтропия	0.8979	0.8903
	Второй угловой момент	0.9032	0.8840
	Несходство	0.8977	0.8812
	Максимум	0.9007	0.8911
	Однородность	0.8964	0.8831
	Энергия	0.8993	0.8901
	Контраст	0.8895	0.8825
Геометрический фильтр Кримминса		0.9153	0.9009
«Нерезкое маскирование»		0.9136	0.9025
Фильтр Габора 1		0.8982	0.8894
Фильтр Габора 2		0.9027	0.8970
Фильтр Габора 3		0.9110	0.9017

Из таблицы 3 видно, что наибольшая точность сегментации для U-Net достигается с использованием геометрического фильтра Кримминса, на втором месте идет фильтр «нерезкое маскирование», затем фильтр Габора №3. Если посмотреть на результаты обучения ResUnet, то можно заметить, что три наилучших результата обучения с дополнительной предварительной обработкой также получены с помощью этих методов.

Наибольшее значение коэффициента Дайса показала сегментация с применением геометрического фильтра Кримминса и архитектурой

ResUnet++ – 0.9175. На рисунке 25 представлены примеры построения масок для данной комбинации.

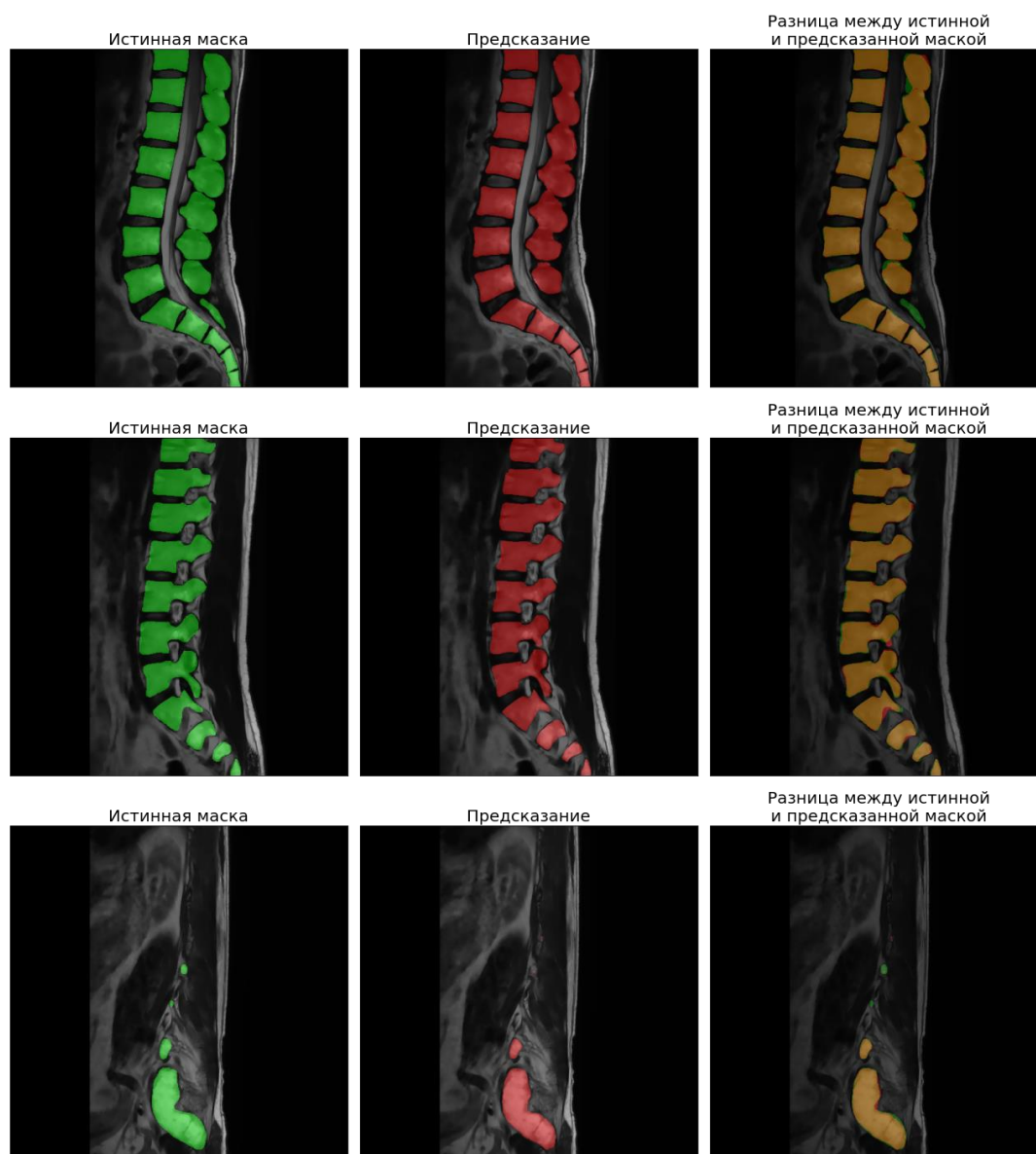


Рисунок 25. Результаты предсказания модели в сравнении с истинными масками. В третьем столбце оранжевым цветом отмечено совпадение истинной и предсказанной масок, зеленым – пиксели маски, которые не обнаружены моделью, красным – пиксели, которые модель неверно присвоила маске

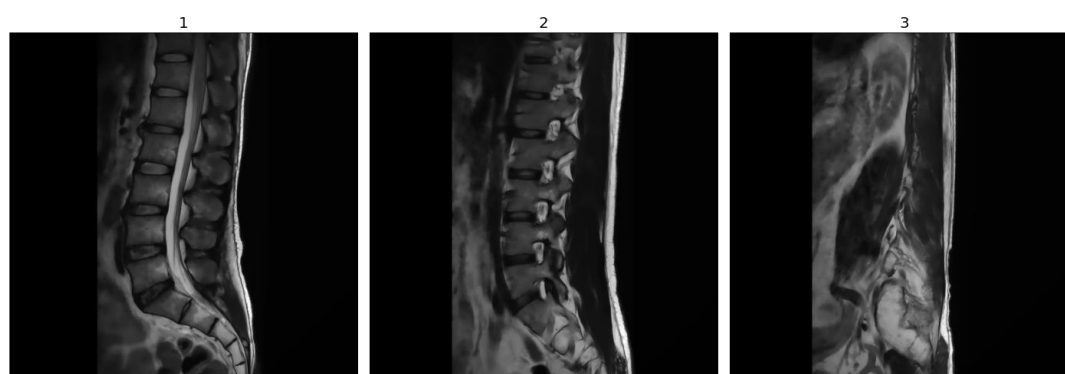


Рисунок 26. Оригинальные изображения для масок с рисунка 25

Выводы

На основе анализа литературы было выбрано несколько методов предварительной обработки изображений: метод, основанный на матрице смежности уровней серого Харалика, фильтры Габора, геометрический фильтр Кримминса и фильтр «нерезкое максирование».

Были реализованы нейронные сети с различными архитектурами: U-Net, ResUnet, ResUnet++, UNet++.

Сравнение результатов обучения без дополнительной предварительной обработки показало, что наибольшее значение коэффициента Дайса имеет нейронная сеть с архитектурой ResUnet++ – 0.9108. Кроме того, для оценки качества были посчитаны значения F1 – меры, которое также показало наибольшее значение при сегментации с использованием данной архитектуры – 0.9091.

Сравнение различных фильтров было проведено для архитектур U-Net и ResUnet. Для обеих архитектур наилучшие значения показали геометрический фильтр Кримминса, фильтр «нерезкое маскирование» и фильтры Габора №3.

Наибольшее значение коэффициента Дайса было получено при использовании модели ResUnet++ совместно с предварительной обработкой геометрическим фильтром Кримминса – 0.9175.

Заключение

В данной работе представлен разработанный алгоритм сегментации МРТ изображений, который включает в себя предобработку, а также применение сверточных нейронных сетей на основе архитектур семейства U-Net.

Предварительная обработка включает в себя: приведение к одному формату и размеру; затемнение областей, не представляющих интереса; для некоторых изображений потребовалась коррекция контрастности и построение маски. Для повышения точности сегментации были предложены несколько методов дополнительной предварительной обработки изображений.

Реализовано несколько моделей сверточных нейронных сетей: U-Net, ResUnet, ResUnet++, UNet++, а также проведено сравнение точности сегментации для различных архитектур и методов предварительной обработки.

Список литературы

1. Cieza, A., K. Causey, K. Kamenov, S.W. Hanson, S. Chatterji et al., 2020. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 369, pp. 2006-2017.
2. Jafarinia, M., T. Vos, S. Lim, M. Naghavi, C. Murray et al., 2020. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 369, pp. 1204-1222.
3. Hartvigsen J, M.J. Hancock, A. Kongsted et al., 2018. What low back pain is and why we need to pay attention. *The Lancet*; 391, pp. 2356–2367.
4. Jusman, Y., L.A. Dewiprabamukti, A.N.N. Chamim, Z. Mohamed, S.N.A.M. Kanafiah and N.H.A. Halim, 2020. Application of Watershed Algorithm and Gray Level Co-Occurrence Matrix in Leukemia Cells Images. 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT), Medan, Indonesia, pp. 9-14.
5. Pratap, T. and P. Kokil, 2019. Approximate Optimization of Gabor Filter Parameters in Application to Blood Vessel Segmentation in Retinal Images. WIE Conference on Electrical and Computer Engineering (WIECON-ECE), IEEE International, pp. 1-5.
6. Wen, Z., R. Feng, J. Liu, Y. Li and S. Ying, 2021. GCSBA-Net: Gabor-Based and Cascade Squeeze Bi-Attention Network for Gland Segmentation. *IEEE Journal of Biomedical and Health Informatics*, 25(4), pp. 1185-1196.
7. Yookwan, W., K. Chinnasarn and B. Jantarakongkul, 2018. Automated Vertebrae Pose Estimation in Low-Radiation Image using Modified Gabor Filter and Ellipse Analysis. 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA), IEEE, pp. 141-146.

8. Cruz-Aceves, I., F. Oloumi, R.M. Rangayyan, J.G. Aviña-Cervantes and A. Hernandez-Aguirre, 2016. Automatic segmentation of coronary arteries using Gabor filters and thresholding based on multiobjective optimization. *Biomedical Signal Processing and Control*, 25: 76-85.
9. Jaware, T.H., V.R. Patil, R.D. Badgujar, S. Bhattacharyya, R. Dey, and R.S. Dhar, 2021. Performance Investigations of Filtering Methods for T1 and T2 Weighted Infant Brain MR Images. *Microsystem Technologies*, pp. 1-13.
10. Loizou, C.P., C.S. Pattichis, C.I. Christodoulou, R.S. Istepanian, M. Pantziaris and A. Nicolaides, 2005. Comparative evaluation of despeckle filtering in ultrasound imaging of the carotid artery. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 52: pp. 1653-1669.
11. Hille G., S. Saalfeld, S. Serowy, and K. Tönnies, 2018. Vertebral body segmentation in wide range clinical routine spine MRI data. *Computer Methods and Programs in Biomedicine*, 155: 93–99
12. Mikulka J., D. Chalupa, J. Svoboda, M. Filipovič, M. Repko and M. Maxová, 2020. Multimodal and Multiparametric Spatial Segmentation of Spine. 19th International Conference on Mechatronics-Mechatronika (ME), IEEE, pp. 1-5.
13. Gawel D., P. Głowka, T. Kotwicki and M. Nowak, 2018. Automatic spine tissue segmentation from MRI data based on cascade of boosted classifiers and active appearance model. *BioMed Research International*, vol. 2018, pp. 13.
14. Jamaludin, A., 2017. Automated analysis of spinal MRI using deep learning, PhD thesis, St Hilda's College, Oxford.
15. Lu J-T., S. Pedemonte, B. Bizzo et al., 2018. Deep spine: automated lumbar vertebral segmentation, disc-level designation, and spinal stenosis grading using deep learning. *Machine Learning for Healthcare Conference*, pp. 403-419.

16. Lessmann N., B. van Ginneken, P.A. de Jong, and I. Išgum, 2019. Iterative Fully Convolutional Neural Networks for Automatic Vertebra Segmentation and Identification. *Medical Image Analysis*, 53 pp. 142–155.
17. Zukić D., A. Vlasák, J. Egger, D. Hořínek, C. Nimsky and A. Kolb, 2014. Robust detection and segmentation for diagnosis of vertebral diseases using routine MR images. *Computer Graphics Forum*, 33(6), pp. 190-204.
18. Haralick R. M. Statistical and structural approaches to texture, 1979. *Proceedings of the IEEE*, 67(5), pp. 786-804.
19. Haralick R. M. K. Shanmugan, I. Dinstein, 1979. Textural Features for Image Classification. *IEEE Transactions on systems, man and cybernetics*, vol. 6, pp. 610-621.
20. Crimmins, T. R. 1986. Geometric filter for reducing speckle. *Optical Engineering*, 25(5), pp. 651–654.
21. Unsharp masking [Электронный ресурс]: Википедия. Свободная энциклопедия. – Режим доступа: https://en.wikipedia.org/wiki/Unsharp_masking (дата обращения: 10.04.2021).
22. Gabor filter [Электронный ресурс]: Википедия. Свободная энциклопедия. – Режим доступа: https://en.wikipedia.org/wiki/Gabor_filter (дата обращения: 5.04.2021).
23. Ronneberger O., P. Fischer and T. Brox, 2015. U-Net: convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, Springer, Cham, pp. 234-241.
24. Zhang, Z., Q. Liu, and Y. Wang, 2018. Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5), pp.749-753.
25. Jha D, P.H. Smedsrud, M.A. Riegler et al., 2019. ResUNet++: An advanced architecture for medical image segmentation. *IEEE International Symposium on Multimedia (ISM)*, pp. 225–2255.

26. Zhao T. et al. UNet++-Based multi-channel speech dereverberation and distant speech recognition, 2021 12th International Symposium on Chinese Spoken Language Processing (ISCSLP), Chinese Spoken Language Processing (ISCSLP), pp. 1–5.