

Санкт–Петербургский государственный университет

Спиридонов Александр Юрьевич

Выпускная квалификационная работа
*Нейросетевые методы анализа изображений
аорты*

Уровень образования: магистратура

Направление: 02.04.02 «Фундаментальная информатика и
информационные технологии

Основная образовательная программа ВМ.5503.2020 «Технологии баз
данных»

Научный руководитель:

Кандидат технических наук, доцент кафедры технологии
программирования - Блеканов Иван Станиславович.

Рецензент:

Начальник отдела стратегических инициатив, Центр разработки и
монетизации данных ООО «Газпромнефть-Цифровые решения» -

Кононов Ярослав Сергеевич.

Санкт-Петербург

2022 г.

Содержание

Введение	3
Обзор литературы	6
Обзор существующих решений	8
Глава 1. Данные	9
1.1. Наборы данных	9
1.2. Формат исходных данных	10
1.3. Подготовка	10
1.3.1 Обработка	10
1.3.2 Объединение просветов	12
1.3.3 Разделение данных	12
1.4. Программная реализация	15
Глава 2. Машинное обучение	16
2.1. Методология исследования	16
2.2. Аугментации	16
2.3. Рассматриваемые модели и архитектуры	18
2.4. Параметры обучения	20
2.5. Результаты	21
2.6. Выводы	22
Глава 3. Измерение диаметра	24
3.1. Введение	24
3.2. Алгоритм	25
3.3. Выводы	25
Заключение	26

Введение

Актуальность

Машинное обучение переживает небывалый рассвет. В виду увеличения вычислительных мощностей и большого объёма информации область машинного обучения получила огромный толчок в развитии. Начиная с 2012 года, когда Джеффри Хинтон представил метод обратного распространения ошибки[36], нейросети начали стремительно набирать популярность. Получив возможность обучать искусственные нейронные сети, компьютеры смогли выполнять задания, бывшие прерогативой людей. В итоге, на сегодняшний день, машинное обучение присутствует во всех аспектах нашей жизни. Человечество разговаривает с умными колонками, передвигается на машинах с автоматическим управлением и много более технологических достижений. Медицина не остаётся в стороне и имеет свои достижения[26] с применением машинного обучения. В данной работе будут рассмотрены последние технологии нейросетевых моделей на примере аорты.

Аорта является примером артерий эластического типа, стенка аорты состоит из трех слоев: внутреннего или интимы (*tunica intima*), среднего или меди (tunicamedia) и наружного или адвентиции (*tunica adventitia*). Основной пагенетической особенностью диссекции аорты является разрыв интимы с последующим образованием и распространением субинтимальной гематомы. Гематома, распространяющаяся вдоль слоев аортальной стенки, обычно занимает от 1/2 до 3/4 ее диаметра, в некоторых случаях расслоение распространяется циркулярно по всей окружности аорты. Вследствие чего формируется «ложный» просвет, или двухпросветная аорта, что существенно увеличивает риски мальперфузии внутренних органов за счет нарушения кровотока по магистральным артериям, отходящим от аорты. При распространении диссекции на восходящий отдел аорты может развиваться тампонада полости перикарда

Аневризма восходящего отдела аорты – заболевание, характеризующееся сочетанием минимальных клинических проявлений либо их отсут-

ствием, и высокой вероятностью драматически быстро наступающих тяжелых, зачастую, фатальных осложнений. Основными осложнениями данной патологии являются расслоение и/или разрыв аорты. Встречаемость аневризм грудного отдела аорты в популяции составляет 0,16–1,15%. Частота развития расслоения грудной аорты в мире в среднем составляет 3-8 случаев на 100.000 населения в год. При этом отмечается крайне высокий риск внезапной смерти. Ориентировочно две трети диссекций аорты происходят в восходящем отделе (типа А по Стенфордской классификации) и лишь в одной трети случаев наблюдается изолированное расслоение нисходящего отдела аорты (тип В).

При острой диссекции аорты типа А (по Stanford) летальность достигает 50% в первые двое суток от начала заболевания или 1-2% в час, и в течение первого месяца от момента расслоения составляет 75-90% без хирургического вмешательства. Огромное значение имеет фактор времени, от момента расслоения до операции. Среди оперированных пациентов 30-дневная госпитальная летальность до 1985 г. была более 30%, а к 2013-2015 гг. составляла уже 11-12,5%, 5-летняя выживаемость в середине 1990-х гг. составляла 50-71%, в середине 2000-х гг. -63-94%, 10-летняя выживаемость -37-64% и 65-84%, соответственно. Также стоит отметить, что при хроническом расслоении 5-летняя выживаемость без операции составляет 10-15%. Необходимость в хирургическом лечении при остром расслоении по абсолютным показаниям обусловлена крайне высоким риском летального исхода при естественном течении патологии и бесперспективности консервативных методов лечения. В то же время, стабильное состояние пациентов при хроническом проксимальном расслоении аорты дает возможность провести полную диагностику и подготовку к плановой операции. Операции по протезированию аорты при диссекции грудной аорты и торакоабдоминального отдела являются одними из самых сложных. Количество осложнений и высокая госпитальная летальность при данной патологии существенно снизились за последние 20 лет, так, при острой диссекции госпитальная летальность снизилась с 50% до 17-24% в среднем (от 9 до 30%), с приходом новых техник хирургии и технологий, однако и по сей день подобные вмешательства являются настоящим «вызовом» для хирурга.

Общепринятые диагностические алгоритмы и хирургические методики на сегодняшний день недостаточно развиты. Для улучшения как ближайших, так и отдалённых результатов лечения пациентов данной категории необходимо дальнейшее изучение данной проблемы.

Таким образом, тема и научно-практическая значимость обсуждаемой проблемы являются весьма актуальными, поскольку многие вопросы диагностики и лечения расслоения аорты остаются до конца не изученными по настоящее время.

Постановка задачи

В рамках выполнения данной работы автор сотрудничал с врачами из клиники высоких медицинских технологий им. Н. И. Пирогова. Результатом данной кооперации стал обмен знаниями о предметных областях. Сотрудники клиники предоставили необходимую информацию о заболеваниях аорты, а также обозначили ряд проблем в диагностике, предоставив примеры данных для изучения. Данные представляют собой двумерные снимки, полученные с помощью компьютерной томографии. Проблемой анализа снимков является изменение поперечного сечения аорты и обнаружение диссекции.

В результате, были сформированы следующие цели и задачи.

Целью данной работы является разработка нейросетевых методов для проведения анализа изображений аорты. Для достижения вышеупомянутой цели в рамках данной работы рассматриваются следующие задачи:

- подготовка датасета;
- сегментация аорты на изображениях;
- сегментация просветов аорты на изображениях;
- определение диаметра аорты на изображениях.

Решения поставленных задач, обусловило следующую структуру работы. Подготовка датасета, предобработка снимков, разделение на тренировочные и валидационные наборы представлены в главе 1 "Данные". Опи-

сание использованных нейросетевых моделей, аугментаций и метрик отображено в главе 2 "Машинное обучение". В третьей главе написано о задаче определения диаметра аорты на изображениях. В заключении сделаны итоговые выводы работы.

Практическая значимость работы

Результаты данной работы имеют значение для диагностики заболеваний аорты с использованием компьютерных технологий для медицинских организаций и, в частности, для клиники высоких медицинских технологий им. Н. И. Пирогова.

Обзор литературы

Проблема изучения аорты исследуется уже не первое десятилетие. До 2015 года работы по анализу аорты и её рассечений опирались на использовании методов обработки изображений. Например, Anderson[2] в своей работе представил полуавтоматический алгоритм сегментации для определения диссекции, основанный на поиске краёв на изображении. Аналогичный подход использовал Krissian[44], создав полуавтоматический алгоритм с изучением контрастных точек, дополнив его размытием мешающих высококонтрастных органов. Помимо анализа краёв, в работе Lee[32] использовался вэйвлет анализ для различия просветов аорты. Полноценные сравнения существующих методов также появились вместе с их развитием[1; 20]. Все вышеописанные методы, представляют собой полуавтоматическую настроенную ручную систему. Подобные системы из-за необходимости в ручной настройке, при работе в новых условиях склонны к ошибкам[20].

Начиная с 2012 года под влиянием работы Hinton начались активные исследования применения нейронных сетей, заинтересовавших людей высокой точностью и автоматичностью[30]. В 2015 году начали появляться первые модели показавшие работоспособность глубоких нейросетей в задачах обработки медицинских снимков[18]. В последнее время, с помощью глубокого обучения выдающиеся результаты были достигнуты почти во всех областях анализа медицинских изображений [3]. В особенности много

усилий было приложено к изучению автоматической сегментации при помощи свёрточных нейросетей (Convolutional Neural Networks, CNN) и сетей с глубоким обучением (Deep Neural Networks, DNN)[33]. Среди них в последние годы появляются новые архитектуры, направленные на улучшение качества сегментации. Например, в работе Pereira S[12] изучен потенциал свёрточных нейронных сетей с малым ядром для сегментации MRI снимков вместе с нормализацией интенсивности.

Нейронные сети используются в обработке изображений аорты. Например, в работе Yu[4] авторы использовали 3D U-net[42] для сегментации аорты и её просветов. Однако, появились более современные модели и в рамках данной работы рассматриваются только двумерные снимки. В работе Akinori[17] использовал предобученный Xception[13] для классификации наличия диссекции аорты на двумерных снимках компьютерной томографии, но без выделения просветов. Таким образом, применение предобученных нейросетей актуально в рамках поставленной проблемы и будет в дальнейшем изучено в работе.

В 2020 году были выпущены две статьи, в которых авторы представили архитектуры, имеющие в себе сразу две объединённые вместе нейросети[5; 34]. В своей работе Li F.[34] опробовал двухпоточную архитектуру для борьбы с ошибками сегментации. Одна нейросеть занималась поверхностными признаками, а вторая наоборот глубокими. Такой дизайн из двух параллельных потоков, согласно автору, показал результаты точности сегментации на уровне передовых моделей. В тоже время, Conze[5] в своей работе использовал также две нейросети, но расположенные не параллельно, а последовательно. Во избежание недообучения результаты первой сети объединяются с исходным изображением перед прохождением сквозь вторую сеть. Данная архитектура, по словам автора, заняла первое место в соревновании CHAOS по сегментации в трёх номинациях. Подобная архитектура использовалась в работе Yu[4], где авторы применили аналогичный подход для выявления расслоения аорты с помощью двух последовательных моделей: для сегментации аорты и просветов соответственно. Однако данная модель, по словам авторов, работает достаточно долго и требует множества снимков на разных уровнях. Результаты данных работ застав-

ляют задуматься о возможности объединения рассмотренных архитектур. Тем не менее, как показывают результаты Yu[4], предварительная сегментация аорты улучшает качество диагностирования, что и будет основным акцентом в дальнейшем исследовании.

Обзор существующих решений

Если сравнивать используемые модели для сегментации, то заметно преобладание U-net[42] архитектуры. Несмотря на появление новых и более сложных моделей, U-net всё ещё активно используется во многих работах, связанных с медициной и показывает достойные результаты[4; 5; 9; 21]. Также популярны свёрточные нейросети без строгой архитектуры[8; 12; 23; 35]. Нужно отметить, что в рамках данной работы будут рассматриваться архитектуры с использованием предобученных моделей. Причиной такого выбора является возможность переноса весов из передовых моделей, без необходимости обучать таковую, что частично нивелирует вычислительные ограничения, не принося в жертву качество.

Необходимо также рассмотреть последние достижения в области работы с изображениями, не ограничиваясь областью медицины. В рамках решения задачи сегментации, важно обратить внимание на статьи, в которых проведено полноценное сравнение существующих моделей и архитектур используемых для сегментации[15; 27]. Авторы отмечают, что U-net уступает DeepLabV3[41] и DeepLabV3Plus[22]. Особенно стоит отметить комбинацию DeepLabV3Plus с Xception71[13], которая показала лучшие результаты среди предобученных моделей. Тем не менее U-net всё ещё рассматривается как базовое решение, относительно которого проводится сравнение моделей. Этот подход, будет применён в дальнейшей работе, где U-net с энкодером ResNet-50[19] будет использоваться как стандарт.

Отдельно заслуживают изучения модели, появившиеся под влиянием статьи "Attention Is All You Need"[10], представившей миру концепт "внимания". Данная работа привела к появлению нового класса моделей, под названием трансформеры. Механизм "внимания" уже показал[31] свою эффективность в области обработки естественного языка, и тем самым его применение в области компьютерного зрения было вопросом време-

ни. Рассматривая предобученные модели с трансформерами, можно выделить несколько важных для данной области статей. Работа Dosovitskiy[7] представляет особый интерес, так как помимо текста авторы практически доказали эффективность трансформеров в области компьютерного зрения, побив несколько рекордов. Помимо этого, авторы ViT выложили веса своей модели в открытый доступ, открыв путь новым архитектурам. Предобученный ViT используется, в свою очередь, в моделях для сегментации, как например в работе Chen[48]. TransUnet представлен авторами как трансформерная сеть, предназначенная для сегментации медицинских снимков, становясь важной моделью для проверки. Также стоит отметить Swin[47]. Swin на данный момент является State-of-the-art решением для сегментации на ADE20K датасете, согласно результатам[37]. Несмотря на то, что Swin не предназначен для работы с медицинскими снимками, как передовое решение, он вызывает интерес в рамках задачи сегментации.

Глава 1. Данные

1.1 Наборы данных

В качестве данных для обучения моделей машинного обучения, были найдены два набора.

Первый набор данных "SegTHOR"[43] содержит в себе 3D снимки компьютерной томографии 40 пациентов. На этих снимках размечены сердце, аорта, трахея и пищевод. В рамках данной работы, использовались данные аорты.

Второй набор "ImageTBAD"[28] содержит в себе также 3D снимки компьютерной томографии, с размеченными просветами аорты. В нём выделено всего три класса: истинный просвет, ложный просвет и тромб ложного просвета. Также стоит отметить, что не все файлы пациентов содержат в себе все виды классов. Как отмечают авторы набора, тромб ложного просвета присутствует у 68 пациентов из 100. Дисбаланс классов недостаточно большой для необходимости принятия специальных мер. Помимо этого, для данного набора удалось получить данные настроек томографа,

а точнее пиксельные интервалы. Настройки томографа влияют на качество снимков и, что важнее, в рамках данной работы, позволяют установить соответствие размеров на изображении к реальным.

1.2 Формат исходных данных

В обоих рассматриваемых наборах, данные предоставлялись в виде zip архивов, содержащих в себе nii.gz 3D файлы компьютерной томографии и соответствующей разметки. На данном этапе было принято решение разделить данные на отдельные срезы, которые будут использованы для обучения. Есть несколько причин для применения подобного подхода, вместо обучения напрямую из 3D файлов. Главной причиной является, появление возможности чередовать по очереди снимки различных пациентов. При обучении с использованием 3D файлов для смешивания снимков необходимо либо держать несколько 3D файлов в памяти сразу, либо часто считывать их по очереди, либо создавать специальные тренировочные 3D файлы для обучения. Такой подход требует больше видеопамати, так как размер файлов в среднем 150Мб для снимков и 5Мб для масок. Это может вызвать проблемы, либо со скоростью, как в случае с частым считыванием, либо с использованной памятью, как в случае с несколькими пациентами сразу. В тоже время без перемешивания файлов модель может переобучиться, из-за получения снимков одного пациента. Вариант, со специальными 3D файлами, не рассматривается по причине неэффективного использования памяти.

1.3 Подготовка

1.3.1 Обработка

Каждый 3D файл содержит от 150 до 300 срезов, в зависимости от настроек томографа. Как показал эксперимент, если брать слишком частые срезы, то модель начнёт быстро переобучаться. Причиной этого является близость между срезами, так как снимки при маленьком изменении высоты, также слабо изменяются. Поэтому для создания обучающего набора

данных был взят каждый 4-ый срез. Данная частота, оказалась достаточно редкой для разнообразия снимков.

Томограф, сделавший снимки, определял границы возможных значений пикселей. В форматах хранения изображений, например, jpeg[29], значения пикселей находятся в промежутке от 0 до 255. Однако в зависимости от настроек томографа промежутков значений может изменяться. В виду данной особенности, перед обучением моделей предварительно требовалось нормализовать цвета, приведя их к единому промежутку значений. После данного этапа снимки сохранялись на диск в формате jpg.

В задачи сегментации требуется выделить объект на снимке. Проходя последовательно по каждому пикселю, модель предсказывает, является ли этот пиксель частью предсказываемого класса. В результате получается новое изображение, где 0 соответствует отсутствию класса, а 1 наличию. Значения наличия и отсутствия классов зависят от функции активации, данный пример соответствует сигмоиде. Однако, в случае сегментирования более чем одного класса, может появиться проблема их наложения. Решением является выделение каждого класса в отдельный канал изображения. Таким образом, модель может решать задачу сегментации для каждого класса последовательно. Тем не менее, это означает что данные для обучения должны также быть созданы с выделенными каналами. Аналогично с проблемой 2D или 3D файлов, для повышения скорости обучения модели создание тренировочных масок было сделано заранее. В итоге, в тренировочных данных получаются пары изображение - набор масок.

Полученный набор масок представляет из себя набор 0 с редкими эллипсами из единиц. Данные изображения представляют их себя разреженные матрицы[46]. Тем не менее, они хранились как полноценные массивы, что было неэффективно. Так как изначально рассматривался вариант обучения моделей на Google Colab[25], требовалось уменьшить размер масок. Для этого маски сохранялись не как изображения или стандартные массивы данных, а как разреженные матрицы. Использование данного метода позволило уменьшить занимаемое место в среднем с 2МБ до 24КБ.

1.3.2 Объединение просветов

Одна из задач данной работы - сегментация аорты на изображениях. Результаты данной сегментации будут использованы для измерения диаметра аорты и сегментации её просветов. Выполнить данную задачу используя только набор данных SegTHOR[43] достаточно сложно. Имея в наличии набор данных ImageTBAD[28], с размеченными просветами, но не аортой, можно сформулировать следующий вопрос. Возможно ли использовать размеченные просветы - части аорты, для сегментации самой аорты? В рамках данной работы был написан алгоритм, использующий стандартные методы обработки изображений для решения данной проблемы.

Идея алгоритма достаточно проста:объединить просветы, минимально покидая зону аорты. Для этого используются морфологические методы обработки изображений: наращивание и эрозия. Предварительно, начальная маска, содержащая метки классов от 0 до 2 бинаризуется. То есть, все ненулевые значения становятся равны единице. После чего, за счёт наращивания, объединяются просветы аорты, заполняя пробелы между просветами. После чего, применяя эрозию, появившиеся излишки за границей аорты убираются.

Таким образом, алгоритм состоит из следующих этапов:

1. Приведение изображения к бинарному, см. рис.1 b.
2. Объединение просветов с помощью наращивания, см. рис.1 c.
3. С помощью эрозии излишки, появившиеся в результате наращивания убираются, см. рис.1 d.

В результате, для сегментации аорты автор использует не один набор данных состоящий из 40 пациентов, а два набора суммарно из 140 пациентов.

1.3.3 Разделение данных

Для успешного обучения модели нужны не только тренировочные данные. В процессе обучения она может начать запоминать увиденные при-

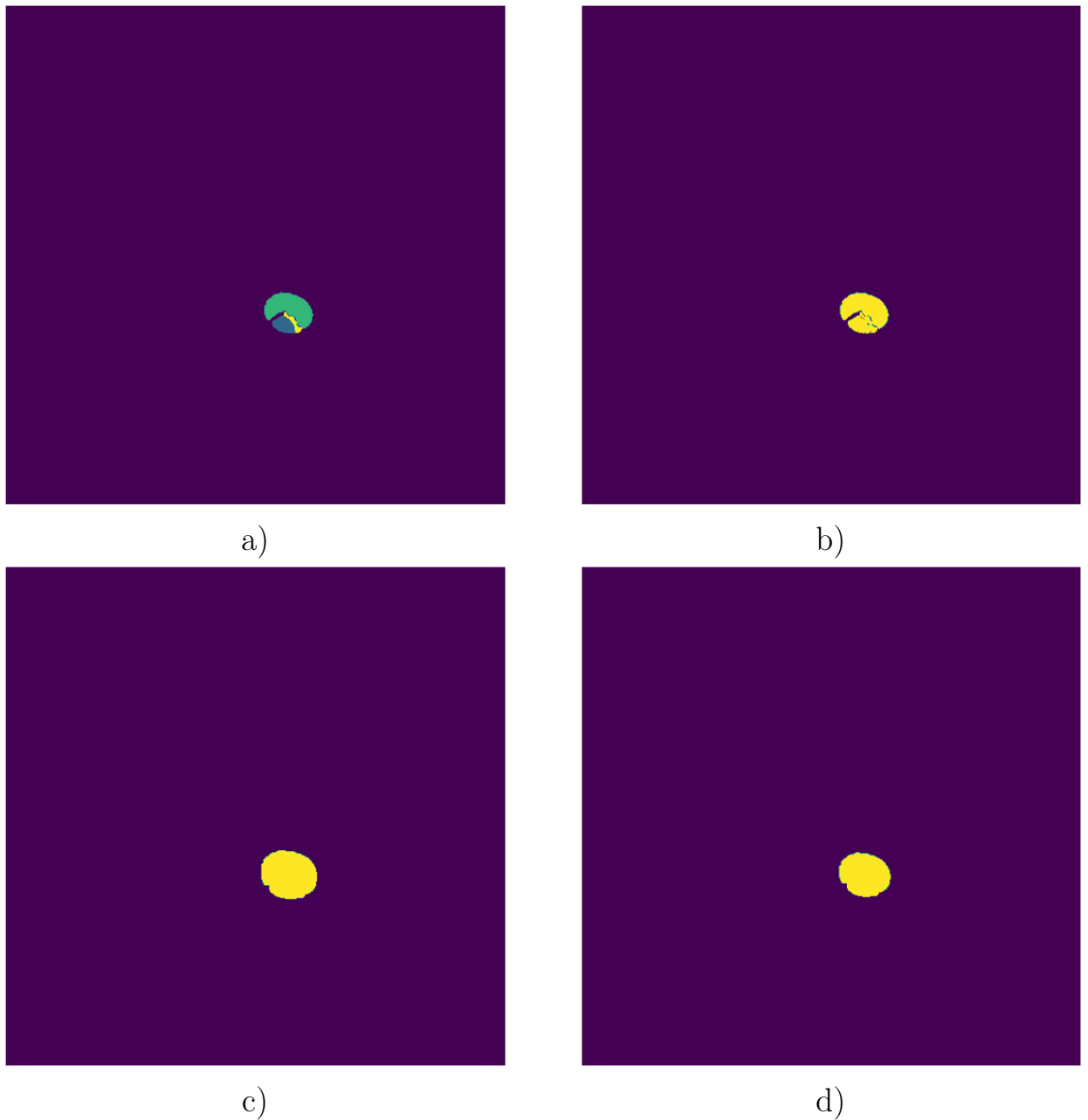


Рис. 1: Объединение просветов: а) изначальные просветы б) бинарное изображение с) результат наращивания д) результат эрозии

меры тренировочных данных, то есть переобучаться. Результатом является модель работающая преимущественно на данных из тренировочного набора. Во избежание переобучения, требуется выделить валидационный (проверочный) набор данных, не используемый в обучении. За счёт проверки модели на валидационном множестве, возможно узнать её состояние.

В современных библиотеках и фреймворках для машинного обучения существуют готовые методы создания валидационной выборки. Принцип

лежащий в их основе достаточно простой - случайно выбрать часть данных, с некой пропорцией. В большинстве случаев, когда отдельно взятые элементы данных не связаны между собой, данный подход успешен. Однако, использовать его с данными используемыми в работе нельзя. Как показала проба обучения, при случайной выборке валидационных данных, происходит утечка тренировочного набора в проверочный. Причиной является схожесть данных от одного пациента, схожесть соседних срезов. Поэтому, при их подготовке автор был вынужден отбросить часть данных, во избежание переобучения. Тем не менее, даже отбросив часть, валидация оказалась "похожа" на тренировочные данные.

По этой причине, для валидации были выбраны не случайные срезы, а отдельные пациенты. Процесс выбора пациентов для валидации был не случаен. Обязательно выбирались "проблемные" пациенты, например, с яркими или тёмными снимками, а также с разнообразной формой тела. В результате, валидационная выборка стремилась быть максимально показательной. Так как количество снимков для каждого отдельного пациента было случайным, при разделении на группы, соотношение валидационных к тренировочным данным выбиралось с погрешностью.

Среди данных также часто выделяется тестовое подмножество данных. Используя модель для предсказания результатов тестового подмножества, возможно получить значения метрик и оценить качество работы модели. Отличие от валидационного множества в том, что его можно использовать для кросс-валидации. Суть кросс-валидации состоит в выборе нескольких пар наборов для тренировки с валидацией и обучение на них модели. В результате появляется возможность сравнения результатов. Однако, кросс-валидация в зависимости от количества разделений пропорционально увеличивает время, требуемое на обучение. Для деления данных на тренировочные-валидационные выборки, потребуется, либо случайно выбирать пациентов из-за чего соотношение изображений будет случайным и не точным, либо вручную разделить данные. С учётом вышесказанного, кросс-валидация в условиях данной работы достаточно проблематична. В остальном, отличий валидационной выборки от тестовой нет. По описанным причинам, тестовая выборка в данной работе не выделялась, а

тестирование проводилось на валидационной выборке.

1.4 Программная реализация

Необходимо отметить наличие разницы в методе хранения данных пациентов. В SegTHOR данные были разделены по папкам, где в каждой папке находится информация касательно одного пациента. В тоже время, в ImageTBAD все файлы снимков и разметки находятся в одной папке. Для подготовки данных автором была написана не одна программа для разметки, а две для каждого набора данных соответственно.

Помимо этого, для объединения наборов данных SegTHOR и ImageTBAD с выделенной аортой был написан код программы. Используя указанные имена файлов пациентов, программа копирует соответствующие файлы в общую папку. Аналогично, данная программа использовалась для создания валидационной выборки.

Все программы были написаны на языке Python 3[49] и будут доступны в репозитории автора[24].

Таким образом, для подготовки файлов были написаны следующие программы:

1. Подготовка SegTHOR;
2. Подготовка ImageTBAD;
3. Подготовка ImageTBAD, с выделением аорты;
4. Перенос файлов выбранных пациентов.

Результатом работы описанных выше программ являются три подготовленных набора данных(см. таблицу 1): набор данных для сегментации аорты содержащий все доступные снимки; набор данных для сегментации просветов аорты, где на входных изображениях выделена аорта; третий набор предназначен для сегментации просветов аорты без изменения снимков.

Таким образом, поставленные во введении задачи выполнены и данные готовы к использованию в алгоритмах машинного обучения.

Набор данных	Тренировочный набор (ед. снимков)	Валидационный набор (ед. снимков)	Процент валидационной выборки от общего числа снимков
Аорта	6524	1483	18,52%
Просветы с аортой	5922	1377	18,87%
Только просветы	5922	1377	18,87%

Таблица 1: Наборы данных

Глава 2. Машинное обучение

2.1 Методология исследования

В рамках данной работы, автор проверяет идею показанную в статьях Yu[4] и Chen[11]. Она состоит в том, чтобы для повышения точности сегментации просветов аорты, перед обработкой снимков, предварительно сегментировать саму аорту. Согласно работам, указанных выше авторов, подобная архитектура приведёт к повышению качества сегментации просветов. Данный подход рационален, так как сегментация аорты уже используется для решения задачи с определением размера. Проверка гипотезы будет проходить для двумерных снимков, в отличии от трёхмерных, используемых в работах данных авторов.

Таким образом, в рамках данной работы будет проведено сравнение не только моделей между собой, а также будет проверена эффективность дополнительного сегментирования аорты перед сегментированием просветов. Для этого, в предыдущей главе, были подготовлены три набора данных. С использованием данных наборов, будут обучены модели машинного обучения. Лучшие среди них, в задачах сегментации аорты и сегментации просветов с выделенной аортой, будут объединены. Результаты комбинированной модели и модели сегментации просветов, без аорты, будут сравниваться.

2.2 Аугментации

Хорошая модель способна выделить признаки сегментируемого объекта на изображениях, несмотря на различные шумы, помехи, сдвиги, повороты и перепады яркости. Это важно в следствии двух причин: наличия помех в снимках и необходимости в обобщении признаков.

В виду различия в приборах для снятия снимков и человеческих ошибок указанные выше проблемы встречаются, также, в медицинских снимках, см. рис2. Поэтому данный фактор остаётся важным в работе с ними.

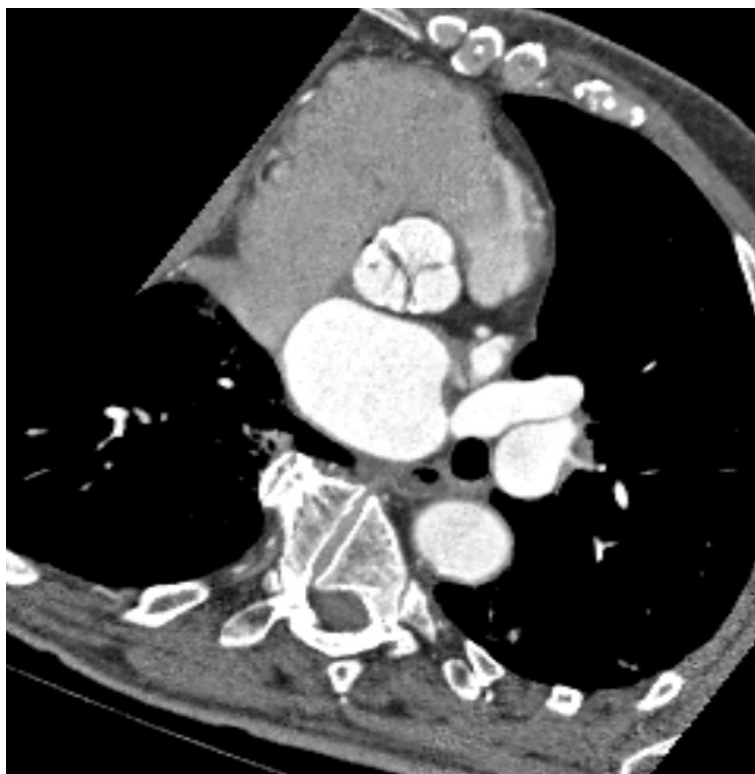


Рис. 2: Пример снимка предоставленный медицинскими работниками клиники высоких медицинских технологий им. Н.И.Пирогова. Снимок находится под углом.

Обобщение важно для всех моделей машинного обучения. Оно подразумевает, что, обучаясь, модель не будет заучивать мелкие детали. Вместо этого, целью обучения модели является выделение характеристик присутствующих только сегментируемому объекту.

Обе вышеуказанные проблемы, можно решить с помощью большого количества данных или аугментаций. Аугментации являются неотъемлемой частью современного машинного обучения. Суть аугментаций состоит в видоизменении изображения в рамках, обозначенных искомой задачей. Например, для задачи распознавания логотипов, вертикальный переворот изображения не поможет, так как ориентация важна для правильного запоминания логотипа.

В рамках задачи сегментации, у аугментаций есть ещё одна важная деталь. В случае с классификацией, нам достаточно применить их к исход-

Аугментация	Параметры	Вероятность
Rotate	(-90;90)	0,6
HorizontalFlip	-	0,5
VerticalFlip	-	0,5
RandomBrightnessContrast	(-0.2;0.2)	0.35
RandomResizedCrop	(256,256)	0.3

Таблица 2: Используемые аугментации.

ному изображению, не учитывая какие конкретно аугментации использовались. При решении задачи сегментации такой подход не подойдёт. Причиной является тот факт, что результат работы модели также является изображением. Соответственно, при изменении исходного изображения, его маска должна измениться аналогично. Исключением же являются аугментации, не влияющие на позицию снимка. Такие аугментации, как например повышение или понижение контрастности, должны применяться исключительно к входным изображениям.

Отдельно стоит отметить, что использование RandomResizedCrop имело и свои минусы. Среди всех аугментаций, данная является самой эффективной в создании "новых" экземпляров данных. Однако, во время использования с некоторыми моделями, применение данной аугментации приводило к ухудшению результатов. Более детально данное наблюдение будет расписано в следующем параграфе работы.

2.3 Рассматриваемые модели и архитектуры

Среди всевозможных моделей и архитектур сегментации, в области машинного обучения, в рамках данной работы будут рассмотрены несколько, выделенных в обзоре литературы.

Во-первых, в качестве базовой, проверенной временем модели был выбран U-net с энкодером ResNet-50. Как уже описывалось в обзоре литературы, несмотря на "возраст" архитектуры U-net, она достаточно часто используется.

Во-вторых, согласно работе Minae[27] одну из лидирующих позиций занимает комбинация DeepLabV3Plus[22] и Xception71[13]. В рассмотрении

предобученных моделей, данная пара заняла первое место. Тем не менее, эффективность данной модели не проверялась для задачи медицинской сегментации. Таким образом, данная комбинация представляет интерес в рамках данной работы.

В-третьих, будет проверена модель с использованием энкодера в основе которого лежит трансформер. С помощью обзора литературы были выделены две таких модели: TransUnet[48] и SWIN[47]. Первоначально планировалось проверить обе модели, используя их из репозитория авторов. Однако, в случае с SWIN из-за сложности применения модели в работе, она не будет использоваться. Модель TransUnet была взята из репозитория автора, однако пайплайн обучения не используется. Причиной является его разительное отличие от кода применяемого для остальных моделей. При обучении моделей с такими различиями в подходе, их сравнение не является осмысленным, так как нельзя однозначно установить источник отличий в метриках. Тем не менее, подобные условия обучения были восстановлены в данной работе, после анализа кода обучения TransUnet. Они включают использование специальной функции потерь и уменьшения темпа обучения линейно, которые подробно описаны в параграфе работы "Параметры обучения". Так как данные детали не изменяют пайплайн коренным образом, для достижения результатов, аналогичных описанным в работе вышеуказанных авторов, функция потерь и темп обучения были перенесены вместе с моделью. Необходимо обратить внимание на то, что в отличие от остальных моделей, описанных в работе, для обучения данной модели не используются RandomResizedCrop. Причиной является архитектура, а точнее преобученная модель, лежащая в её основе. TransUnet использует в работе ViT[7], построенный на основе текстовой модели. Одна из основных идей ViT, это разделение изображения на 16 частей, каждая из которых подаётся трансформеру, см рис.3. Автор предполагает, что использование трансформеров с подобным разделением, позволяет модели учитывать информацию со всего изображения. Как показало пробное обучение, данная гипотеза подтвердилась. Попытка обучения остальных моделей без RandomResizedCrop приводила к переобучению, в отличие от TransUnet.

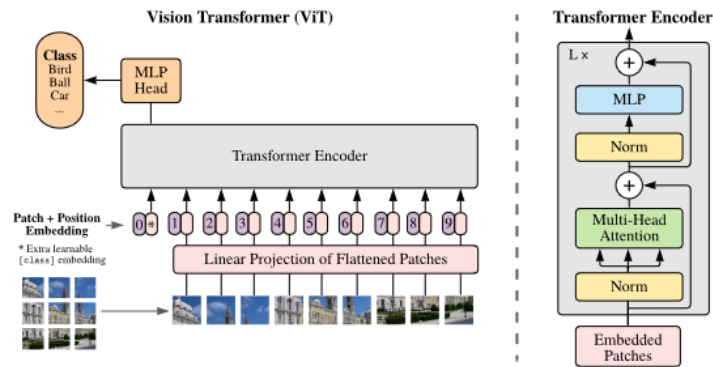


Рис. 3: Пример разделения изображения в ViT архитектуре. Источник:[7]

2.4 Параметры обучения

Основной язык, который использовался автором для программирования алгоритмов машинного обучения, Python 3[49]. В качестве фреймворка был выбран PyTorch[40], в виду использования его в передовых моделях и соответственно возможности использовать в работе TransUnet[48].

Возможности применения одинаковых аугментаций к изображениям и маскам в PyTorch на данный момент нет. Данную задачу выполняла библиотека Albumentations[6].

Для построения пайплайна обучения использовалась библиотека Segmentation Models Pytorch(SMP)[51]. Она была выбрана вследствие легкости и удобства замены моделей и архитектур. Надёжность библиотеки также подтверждается наличием соревнований Kaggle выигранных с её использованием, см. репозиторий библиотеки[51].

Последней важной библиотекой, использованной в данной работе был torch-lr-finder[39]. Она реализует алгоритм, предложенный в работе Smith[45], подбора темпа обучения. Её преимущество состоит в построении графика изменения функции потерь в зависимости от темпа обучения. Указанный метод, позволяет уменьшить случайность в определении темпа обучения.

Показатель темпов обучения в процессе обучения изменялся при помощи learning rate schedulers(lrs). С использованием lrs темп обучения меняется в зависимости от номера эпохи, значений функции потерь и т.п. Существует множество различных видов lrs[38], в рамках данной работы используются два вида. Первый, lrs меняет темп "шагами" в зависимости

от номера эпохи и используется для U-net и DeepLabV3Plus. Данный lrs является стандартным для задач машинного обучения и выбран в виду простоты и практичности. Второй, lrs меняет темп линейно в зависимости от номера эпохи и используется для TransUnet. Этот lrs был выбран, так как именно он использовался при обучении TransUnet, описанном в соответствующей статье.

$$DL(y, \hat{y}) = 1 - \frac{2y\hat{y} + 1}{y + \hat{y} + 1} \quad (1)$$

В качестве функции потерь изначально была выбрана Dice loss, см. формулу 1. Однако, впоследствии была использована функция потерь аналогичная используемой в работе про TransUnet[48]. Новая функция потерь, называемая Combo[14], при проведении тестирования показала улучшение всех основных метрик. По своей идее она представляет линейную комбинацию Dice loss и Binary Cross-entropy loss[16], см. формулу 2.

$$CL(y, \hat{y}) = \frac{1}{2} * (DL(y, \hat{y}) + BCE(y, \hat{y})),$$

$$\text{where } BCE(y, \hat{y}) = -(y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})) \quad (2)$$

В качестве машины для обучения изначально рассматривался Google Colab[25]. Однако в виду ограничений на время непрерывной работы, обучение было перенесено на личную машину автора. Таким образом, обучение проходило на GTX 1060 6gb. Все модели обучались 30 эпох с ранней остановкой в случае отсутствия изменений на протяжении 5 эпох.

2.5 Результаты

Для оценки результатов работы модели будут рассматриваться следующие метрики.

Intersection over union(IoU), используется для оценки качества моделей сегментации и детектирования объектов. В основе IoU лежит коэффициент Жаккара, см. формулу 3.

$$IoU(A, B) = J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (3)$$

Метрика Recall, в рамках данной задачи показывающая отношение правильно сегментированных частей аорты/просветов (True Positive) к всей части снимка которая должна была быть выделена, см. формулу 4.

Метрика, F1 (см. формулу 5) которая также равняется коэффициенту Сёренсена (Dice similarity coefficient) и отображает Dice loss.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$F_1 = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = DSC = \frac{2|X \cap Y|}{|X| + |Y|} \quad (5)$$

Результаты сегментации будут представлены в таблицах 3 - 5. "Жирным" подсвечены лучшие значения метрик среди всех моделей, в одной задаче. Результаты использования двух последовательных моделей, для сегментации аорты и просветов, представлены в таблице 5, для сравнения результатов с вариантами специализированных моделей.

Название	IoU	Recall	F1-score
U-net + Resnet50	0.8632	0.9124	0.9263
DeepLabV3Plus + Xception71	0.8732	0.9236	0.9319
TransUnet	0.8951	0.9182	0.9444

Таблица 3: Значения метрик моделей сегментации аорты.

Название	IoU	Recall	F1-score
U-net + Resnet50	0.5675	0.7388	0.7214
DeepLabV3Plus + Xception71	0.7558	0.8522	0.8577
TransUnet	0.6724	0.7708	0.8004

Таблица 4: Значения метрик моделей сегментации просветов, с выделенной аортой.

2.6 Выводы

Результаты, показанные в предыдущем параграфе, позволяют сделать несколько выводов.

Название	IoU	Recall	F1-score
U-net + Resnet50	0,4233	0.5855	0.5927
DeepLabV3Plus + Xception71	0,7213	0,8093	0,8347
TransUnet	0,6628	0,7479	0,7931
Последовательные TransUnet и DeepLabV3Plus	0.1905	0.4667	0.3099

Таблица 5: Значения метрик моделей сегментации просветов, без выделенной аортой.

Во-первых, лучше всего в задаче сегментации аорты показал себя TransUnet. Использование в качестве энкодера Visual Transformer (ViT), показывает лучшие результаты среди проверенных моделей. Как показано в последующих задачах, TransUnet уступает по показателям DeepLabV3Plus вместе с Xception71. Предположительно, в более простых задачах, таких как сегментация аорты, более продвинутый энкодер TransUnet, позволяет вырваться вперед. Тем не менее, в более сложных задачах, TransUnet начинает терять позиции из-за более простой архитектуры U-net, в отличие от DeepLabV3Plus вместе с Xception71. Предположительно Atrous Spatial Pyramid Pooling (ASPP), представленный в работе DeepLabV3Plus[22] улучшил результаты. ASPP позволили слоям декодера учесть больше информации от изначального изображения и вместе с результатами энкодера сделать более точные предсказания. TransUnet в более сложных задачах сумел показать себя лучше ResNet-50 с U-net, более чем на 10 процентов IoU. Таким образом, автор предполагает использование трансформеров и механизма "внимания" в более сложных архитектурах и соответственно значительного улучшения результатов.

Во-вторых, использование модели для предварительной сегментации аорты, показало спорные результаты. Можно заметить, что в задаче сегментации просветов, все модели показали себя лучше на датасете с предвыделенной аортой, см. таблицу 6. Как демонстрируют результаты комбинированной модели, ошибки сегментации аорты нивелируют улучшения. Для улучшения показателей, модель, используемая при предварительной сегментации должна работать с почти идеальной точностью. В итоге, для задачи сегментации просветов, будет выбрана модель для работы с изначальными снимками, чьи результаты показаны в таблице 5.

Метрика	IoU	Recall	F1
Изменение	$6.28 \pm 1,03$	7.30 ± 0.99	5.30 ± 0.87

Таблица 6: Улучшения метрик при выделенной аорте, значения в % \pm двойная дисперсия.

Глава 3. Измерение диаметра

3.1 Введение

При выполнении данной работы, автор нашёл набор данных "ImageTBAD" с данными томографа. Для решения поставленной задачи определения размеров аорты методами машинного обучения требуются размеченные данные. В рамках данной задачи, подобные данные бы представляли из себя набор изображений и соответствующих расстояний. Имея набор данных с известными пиксельными интервалами возможно получить такой датасет. Для этого есть два варианта: ручной и автоматический. Ручной вариант, подразумевает измерение расстояний руками на каждом снимке. Данный подход чрезвычайно кропотлив, и не спасает от ошибок, и в зависимости от задачи может требовать соответствующего образования. Автоматический подход, с другой стороны быстрый, но ограничен техническими возможностями и более склонен к ошибкам.

В рамках данной работы, будет использоваться автоматический подход. При подготовке данных для машинного обучения, точность является важным фактором, так как модель может запоминать ошибки разметки и начать допускать их сама. Если алгоритм разметки сможет точно подготовить данные, то необходимость в обучении модели отпадает. Если, наоборот, алгоритм будет плохо подготавливать данные, то модель обучить нельзя из-за ошибок. Поэтому, созданный алгоритм будет использоваться непосредственно в задаче измерения размеров аорты.

Таким образом, было решено создать рекомендательную систему. В основе системы лежит алгоритм разметки и результат сегментации аорты. Результатом является примерная оценка размеров аорты. В дополнение должна появиться возможность графического отображения процесса работы, для верификации результата.

3.2 Алгоритм

При изучении снимков аорты можно заметить, что в срезе она представляет фигуру в виде эллипса. Следовательно, на снимках с сегментированной аортой, будет выделена та же фигура. В рамках поставленной задачи требуется найти максимальную и минимальную ширину эллипса, то есть большую и меньшую его полуоси. Полуоси являются параметрами данной фигуры, что приводит нас к задаче нахождения её параметров на изображении. Для этого используется преобразование Хафа, а точнее его улучшенной версией, представленной в работе Yonghong[50].

Отдельно стоит отметить случай наличия на снимке двух аорт. Данный случай возможен из-за того, что форма аорты в профиль представляет собой "крюк". Для подобных снимков будет последовательно проведён анализ сначала одной аорты, потом другой. Смена проводится посредством "закрашивания" одной из аорт.

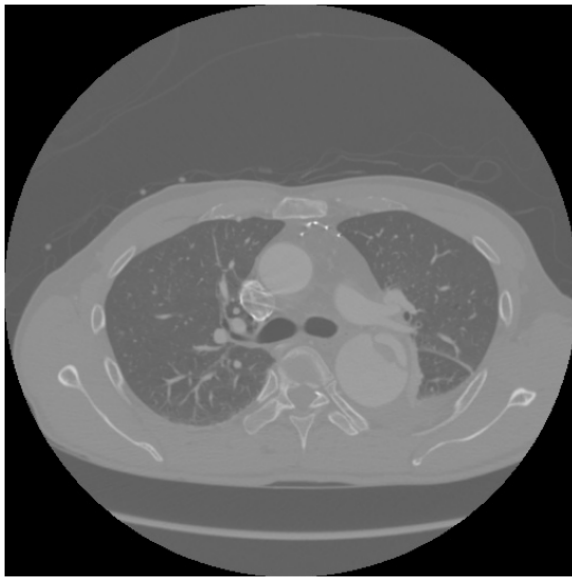
Таким образом, алгоритм состоит из следующих этапов:

1. Выделение краевых точек.
2. Нахождение всех эллипсов и их количества.
3. Определение их характеристик и отображение найденного эллипса.

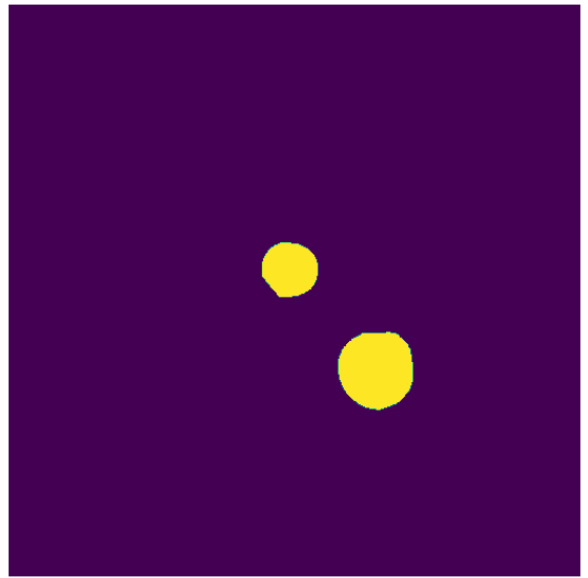
Пример работы алгоритма, включая сегментацию отображён на рисунке 4. Красным нарисован эллипс, соответствующий найденным параметрам. Зелёным нарисована большая ось, а синим малая.

3.3 Выводы

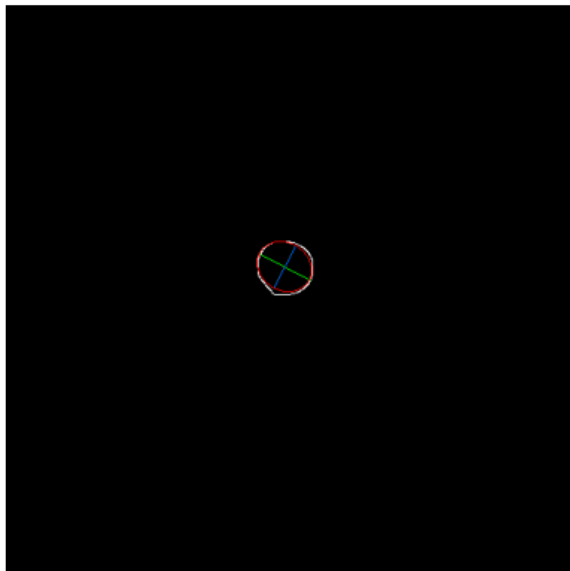
Как можно видеть по рисунку 4, предложенный алгоритм справляется не идеально. Как и предполагалось подобные измерения недостаточно точные для создания датасета. Есть возможность настроить параметры для нахождения оптимального эллипса, однако нет гарантий что алгоритм сработает на другом снимке. Для полноценной оценки, хотелось бы имея размеченный датасет оценить результаты работы. Однако, в виду отсутствия такового, заявлять, что алгоритм точен не представляется возмож-



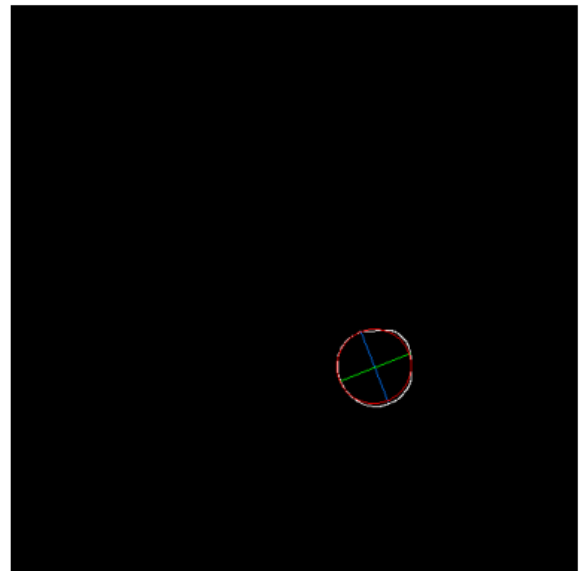
a)



b)



c)



d)

Рис. 4: Измерение аорты: а) изначальный снимок б) результат сегментации с) найденная первая аорта d) найденная вторая аорта

ным. В виду этого, данный алгоритм подходит в качестве рекомендательного. Это означает, что потребуется человек для изучения результатов работы. Тем не менее алгоритм способен найти размеры аорты, в результате чего поставленная задача выполнена.

Заключение

В рамках данной работы была предпринята попытка автоматическо-

го диагностирования расслоения аорты. Для достижения данной цели, автор решил задачу сбора данных и нашёл размеченные снимки компьютерной томографии. В результате их подготовки, был успешно реализован алгоритм для расширения набора данных сегментации аорты. Разработанный алгоритм позволил значительно повысить качество сегментации аорты и избежать переобучения на небольшом датасете. В итоге, были успешно созданы 3 набора готовых к обучению моделей сегментации, которые позволили не только решить поставленные задачи, но и проверить гипотезы, представленные в современных зарубежных работах по этой теме.

Одной из рассмотренных гипотез является использование дополнительной модели машинного обучения для сегментации аорты. По предположению автора, такое 'выделение' объекта должно было локализовать сегментацию просветов аорты, повысив её точность. Было подтверждено улучшение качества сегментации, с помощью проведённого сравнения с использованием 3 различных архитектур, в том числе передовых трансформеров. Однако, также были выявлены негативные эффекты использования дополнительной модели. Ошибки первой модели, значительно понижают качество конечной сегментации. Таким образом две последовательные нейросети показывают результаты значительно хуже, чем использование одной.

В итоге работы с моделями машинного обучения в рамках поставленных целей и задач исследования, автор обучил суммарно 9 нейросетей. Для 3 различных задач, были проверены классический U-net с Resnet50, DeepLabV3 с Xception71 и передовой TransUnet. Результаты, полученные автором, позволили оценить достижения новых моделей и наметить будущие тенденции для улучшения архитектур сегментации. Непосредственным продуктом работы являются обученные передовые модели для сегментации аорты и её просветов. Была достигнута задача обнаружения диссекции и была подготовлена почва для измерения диаметра аорты.

Измерение расстояний на изображениях сложная задача ввиду неизвестности отношения к реальным размерам. Для её решения автором был предложен алгоритм автоматической оценки размеров. Используя данные настроек томографа и результаты сегментации аорты, алгоритм оценивает

минимальный и максимальный диаметр аорты. Данные измерения, в свою очередь, также могут использоваться медиками в работе. Таким образом, последняя поставленная автором задача выполнена.

Данная тема имеет большие перспективы для дальнейшего исследования. Во-первых, результаты данной работы, возможно улучшить, добавив в обучение больше размеченных данных пациентов. Во-вторых, используя дополнительные вычислительные мощности, можно ускорить вычисления и обучить с нуля передовые нейросетевые модели. В-третьих, в будущем возможно использование более сложных моделей сегментации на основе трансформеров. С точки зрения продолжения работы над данной темой, сегментация 3D снимков, выпрямление аорты и классификация типа диссекции только часть возможных путей развития.

Список литературы

1. A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes / D. Lesage [и др.] // *Medical image analysis*. — 2009. — Т. 13, № 6. — С. 819—845.
2. A semi-automated image segmentation approach for computational fluid dynamics studies of aortic dissection / J. R. Anderson [и др.] // *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. — IEEE. 2014. — С. 4727—4730.
3. A survey on deep learning in medical image analysis / G. Litjens [и др.]. — дек.2017.
4. A Three-Dimensional Deep Convolutional Neural Network for Automatic Segmentation and Diameter Measurement of Type B Aortic Dissection / Y. Yu [и др.]. — 2021.
5. Abdominal multi-organ segmentation with cascaded convolutional and adversarial deep networks / P.-H. Conze [и др.]. — 2020.
6. Alumentations: Fast and Flexible Image Augmentations / A. Buslaev [и др.] // *Information*. — 2020. — Т. 11, № 2.

7. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale / A. Dosovitskiy [и др.]. — 2020.
8. Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation / O. Oktay [и др.]. — фев.2018.
9. Artificial Intelligence to Detect Papilledema from Ocular Fundus Photographs / D. Milea [и др.]. — апр.2020.
10. Attention Is All You Need / A. Vaswani [и др.]. — 2017.
11. Automatic detection of aortic dissection in contrast-enhanced CT. In: Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)./Dehghan E, Wang H, Syeda-Mahmood T./ Piscataway, NJ: Institute of Electrical and Electronics Engineers, 2017; 557–560.
12. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images / S. Pereira [и др.]. — май.2016.
13. *Chollet F.* Xception: Deep Learning with Depthwise Separable Convolutions. — 2016.
14. Combo loss: Handling input and output imbalance in multi-organ segmentation / S. A. Taghanaki [и др.] // Computerized Medical Imaging and Graphics. — 2019. — Т. 75. — С. 24–33.
15. Comparison of Deep-Learning-Based Segmentation Models: Using Top View Person Images / I. Ahmed [и др.]. — 2020.
16. *Yi-de M., Qing L., Zhi-Bai Q.* Automated image segmentation using improved PCNN model based on cross-entropy // Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004. — IEEE. 2004. — С. 743–746.
17. Deep learning algorithm for detection of aortic dissection on non-contrast-enhanced CT / A. Hata [и др.]. — авг.2020.
18. Deep learning with non-medical training used for chest pathology identification / Y. Bar [и др.] // Medical Imaging 2015: Computer-Aided Diagnosis. Т. 9414. — International Society for Optics, Photonics. 2015. — С. 94140V.
19. Deep Residual Learning for Image Recognition / K. He [и др.]. — 2015.

20. Detection, segmentation, simulation and visualization of aortic dissections: A review / А. Pepe [и др.] // Medical image analysis. — 2020. — Т. 65. — С. 101773.
21. Diagnostic Accuracy of a Machine-Learning Approach to Coronary Computed Tomographic Angiography–Based Fractional Flow Reserve / А. Coenen [и др.]. — июн.2018.
22. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation / L.-C. Chen [и др.]. — 2018.
23. Full left ventricle quantification via deep multitask relationships learning / W. Хие [и др.]. — янв.2018.
24. Github репозиторий с кодом работы. — 2022. — <https://github.com/WhiteSpirt25/NeuralAortaDiagnostics>.
25. Google Colaboratory. — 2022. — <https://colab.research.google.com>.
26. How is artificial intelligence used in medicine? — 2020. — URL: <https://www.ibm.com/topics/artificial-intelligence-medicine>.
27. Image Segmentation Using Deep Learning: A Survey / S. Minaee [и др.]. — 2020.
28. ImageTBAD: A 3D Computed Tomography Angiography Image Dataset for Automatic Segmentation of Type-B Aortic Dissection / Z. Yao [и др.]. — 2021.
29. JPEG. — 05.2022. — URL: <https://en.wikipedia.org/wiki/JPEG>.
30. *Krizhevsky A., Sutskever I., Hinton G. E.* Imagenet classification with deep convolutional neural networks // Advances in neural information processing systems. — 2012. — Т. 25.
31. Language Models are Few-Shot Learners / Т. В. Brown [и др.]. — 2020.
32. *Lee N., Tek H., Laine A. F.* True-false lumen segmentation of aortic dissection using multi-scale wavelet analysis and generative-discriminative model matching // Medical Imaging 2008: Computer-Aided Diagnosis. Т. 6915. — International Society for Optics, Photonics. 2008. — С. 69152V.

33. *Long J., Shelhamer E., Darrell T.* Fully convolutional networks for semantic segmentation // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2015. — С. 3431—3440.
34. MDFA-Net: Multiscale dual-path feature aggregation network for cardiac segmentation on multi-sequence cardiac MR / F. Li [и др.]. — мар.2021.
35. Multi-stage learning for segmentation of aortic dissections using a prior aortic anatomy simplification / D. Chen [и др.]. — апр.2021.
36. *Pal R.* Deep Learning: The Fastest Growing Tech. in the World. — 2020. — URL: <https://rajtilak1s2510.medium.com/deep-learning-the-fastest-growing-tech-in-the-world-6cd2636a68bf>.
37. Papers with code - the latest in machine learning. — URL: <https://paperswithcode.com/>.
38. Pytorch documentation - optimization algorithms and learning rate schedulers. — URL: <https://pytorch.org/docs/stable/optim.html>.
39. PyTorch learning rate finder. — 2020. — <https://github.com/davidtvs/pytorch-lr-finder>.
40. PyTorch: An Imperative Style, High-Performance Deep Learning Library / A. Paszke [и др.] // Advances in Neural Information Processing Systems 32. — Curran Associates, Inc., 2019. — С. 8024—8035.
41. Rethinking Atrous Convolution for Semantic Image Segmentation / L.-C. Chen [и др.]. — 2017.
42. *Ronneberger O., Fischer P., Brox T.* U-Net: Convolutional Networks for Biomedical Image Segmentation. — 2015.
43. SegTHOR: Segmentation of Thoracic Organs at Risk in CT images / Z. Lambert [и др.]. — 2019.
44. Semi-automatic segmentation and detection of aorta dissection wall in MDCT angiography / K. Krissian [и др.]. — янв.2014.
45. *Smith L. N.* Cyclical learning rates for training neural networks // 2017 IEEE winter conference on applications of computer vision (WACV). — IEEE. 2017. — С. 464—472.

46. Sparse matrix. — 03.2022. — URL: https://en.wikipedia.org/wiki/Sparse_matrix.
47. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows / Z. Liu [и др.]. — 2021.
48. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation J. Chen [и др.] // arXiv preprint arXiv:2102.04306. — 2021.
49. *Van Rossum G., Drake F. L.* Python 3 Reference Manual. — Scotts Valley, CA : CreateSpace, 2009.
50. *Xie Y., Ji Q.* A new efficient ellipse detection method // Object recognition supported by user interaction for service robots. — 2002. — Т. 2. — 957—960 vol.2.
51. *Yakubovskiy P.* Segmentation Models Pytorch. — 2020. — https://github.com/qubvel/segmentation_models.pytorch.