Министерство науки и высшего образования Российской Федерации

Федеральное государственное автономное образовательное учреждение высшего образования

«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина» Институт радиоэлектроники и информационных технологий - РТФ

Кафедра информационных технологий и систем управления

## ДОПУСТИТЬ К ЗАЩИТЕ ПЕРЕД ГЭК

Зав, кафедрой \_\_ИТиСУ Е.В. Кислицын (Ф.И.О.) (прапись) 2024 г. 5 » 01

### ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

МУЛЬТИКЛАССОВАЯ СЕГМЕНТАЦИЯ РАННЕЙ И ПРОМЕЖУТОЧНОЙ ФОРМ ВОЗРАСТНОЙ МАКУЛЯРНОЙ ДЕГЕНЕРАЦИИ НА СНИМКАХ ОПТИЧЕСКОЙ КОГЕРЕНТНОЙ ТОМОГРАФИИ С ИСПОЛЬЗОВАНИЕМ ГЛУБОКОГО ОБУЧЕНИЯ

Научный руководитель: Борисов Василий Ильич к.т.н., доцент

подпись

Нормоконтролер: Бредихина Наталья Сергеевна

подпись

Студент группы: РИМ-220906 Красильникова Юлия Сергеевна

подпись

Екатеринбург 2024 Министерство науки и высшего образования Российской Федерации Федеральное государственное автономное образовательное учреждение высшего образования

## «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

Институт радиоэлектроники и информационных технологий - РТФ Кафедра информационных технологий и систем управления Направление подготовки 09.04.01 Информатика и вычислительная техника Образовательная программа Инженерия искусственного интеллекта

### ЗАДАНИЕ

на выполнение выпускной квалификационной работы

студента <u>Красильниковой Юлии Сергеевны</u> группы <u>РИМ-220906</u> (фамилия, имя, отчество)

1. Тема выпускной квалификационной работы <u>Мультиклассовая сегментация ранней и</u> <u>промежуточной форм возрастной макулярной дегенерации на снимках оптической</u> когерентной томографии с использованием глубокого обучения

Утверждена распоряжением по институту от «4» декабря 2023 г. № 33.02-05/298

2. Научный руководитель Борисов Василий Ильич, доцент, кандидат технических наук (Ф.И.О., должность, ученая степень, ученое звание)

3. Исходные данные к работе набор данных OCTDL (Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods), содержащий снимки оптической когерентной томографии сетчатки глаза

4. Перечень демонстрационных материалов пояснительная записка к ВКР, презентация, выполненная в Microsoft Power Point

#### 5. Календарный план

№ п/п	Наименование этапов выполнения работы	Срок выполнения этапов работы	Отметка о
1.	Глава 1. Исследование алгоритмов нейронных сетей для мультиклассовой сегментации ВМД на снимках ОКТ	до 23.03.2024 г.	Ф
2.	Глава 2. Архитектура нейронной сети UNet3+ для мультиклассовой сегментации ВМД на снимках ОКТ	до 29.04.2024 г.	\$
3.	Глава 3. Обучение модели глубокой сверточной нейронной сети UNet3+ для решения задачи мультиклассовой сегментации ранней и промежуточной форм ВМД	до 19.05.2024 г.	Å
4.	ВКР в целом	до 20.05.2024 г.	A

Научный руководитель <u>Борисов В.И.</u> Ф.И.О.

Студент задание принял к исполнению 04.12.23

дата

6. Допустить <u>Красильникову Юлию Сергеевну</u> к защите выпускной квалификационной работы в экзаменационной комиссии

Зав. кафедрой ИТиСУ

(подпись)

<u>Е.В. Кислицын</u> Ф.И.О.

(подпись)

#### РЕФЕРАТ

Пояснительная записка с. 70, разделов 3, рис. 21, табл. 4, формул 5, источников 60, приложение 1.

СЕГМЕНТАЦИЯ, ГЛУБОКОЕ ОБУЧЕНИЕ, НЕЙРОННЫЕ СЕТИ, ОПТИЧЕСКАЯ КОГЕРЕНТНАЯ ТОМОГРАФИЯ, ОФТАЛЬМОЛОГИЯ

Данная работа посвящена одной из актуальных на сегодняшний день проблем – автоматизации диагностики офтальмологических заболеваний на снимках оптической когерентной томографии (ОКТ). Известно, что автоматизация диагностики решает ряд существующих проблем и сокращает время оценки и интерпретации снимков ОКТ, а также устраняет субъективность оценки, тем самым повышая качество диагностики.

Объектом исследования являются снимки оптической когерентной томографии глаза с ранней и промежуточной формами возрастной макулярной дегенерации.

Предмет исследования разработка алгоритма глубокого обучения для мультиклассовой сегментации «сухой» формы возрастной макулярной дегенерации.

Цель работы – разработать и протестировать модель нейронной сети, способную проводить мультиклассовую сегментацию ранней и промежуточной форм возрастной макулярной дегенерации на снимках оптической когерентной томографии глаза.

Научная новизна: в рамках работы впервые проведена мультиклассовая сегментация «сухой» формы ВМД на открытом наборе данных ОСТDL с использованием сверточной глубокой нейронной сети UNet3+. Архитектура нейронной сети, решающая задачу мультиклассовой сегментации, была разработана на основе сверточной глубокой нейронной сети UNet3+ и обучена на открытом наборе данных ОСТDL.

Практическая значимость работы заключается в том, что использование методов глубокого обучения для сегментации снимков ОКТ и классификации

форм ВМД позволит диагностировать заболевание на ранней стадии и предотвратить его переход в более тяжелую форму. Также использование глубоких нейронных сетей сократит время обследования пациента врачом и снизит влияние субъективной оценки специалиста на постановку диагноза.

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ
1 Исследование алгоритмов нейронных сетей для мультиклассовой
сегментации ВМД на снимках ОКТ 10
1.1 Общая характеристика предметной области 10
1.2 Анализ современных алгоритмов нейронных сетей для сегментации и
классификации ВМД на снимках ОКТ 12
1.3 Биомедицинские аспекты, связанные со строением сетчатки глаза,
офтальмологическими заболеваниями и их диагностикой 17
1.3.1 Строение сетчатки глаза17
1.3.2 Возрастная макулярная дегенерация 20
1.3.3 Гипертоническая ретинопатия
1.3.4 Диабетическая ретинопатия23
1.3.5 Окклюзия центральной вены сетчатки и ветвей центральной вены
сетчатки
1.3.6 Аппаратные методы диагностики заболеваний сетчатки глаза 24
2 Архитектура нейронной сети UNet3+ для мультиклассовой сегментации
ВМД на снимках ОКТ
2.1 Выбор подхода для решения задачи мультиклассовой сегментации 26
2.2 Сверточные нейронные сети
2.3 Алгоритм работы нейронной сети UNet и её модификации
2.4 Метрики оценки качества работы модели нейронной сети
2.5 Выбор инструментов для разработки 38
2.5.1 Выбор программного стека для обучения нейронной сети UNet3+38
2.5.2 Выбор обучающей платформы 40
2.6 Набор данных для обучения 42
2.6.1 Анализ обучающих наборов данных 42
2.6.2 Разметка обучающих наборов данных 46

3 Обучение модели глубокой сверточной нейронной сети UNet3+ для	
решения задачи мультиклассовой сегментации ранней и промежуточной	
форм ВМД	9
3.1 Проведение дополнительной предобработки размеченных	
изображений4	9
3.1.1 Проведение дополнительной предобработки размеченных	
изображений с помощью функции cv2.resize	0
3.1.2 Проведение дополнительной предобработки размеченных	
изображений с помощью Padding5	1
3.2 Настройка и обучение нейронной сети UNet3+	2
3.3 Решение задачи мультиклассовой сегментации ранней, промежуточной	í
и поздней формы ВМД на снимках ОКТ5	5
ЗАКЛЮЧЕНИЕ	8
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 6	1
ПРИЛОЖЕНИЕ А (справочное) Примеры работы модели UNet3+ 6	8

## ПЕРЕЧЕНЬ СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

В настоящем документе применяют следующие сокращения и обозначения:

АФК	– Активные формы кислорода
ВМД	– Возрастная макулярная дегенарация
BO3	– Всемирная Организация Здравоохранения
ГР	– Гипертоническая ретинопатия
ДЗН	– Диск зрительного нерва
ДИ	– Доверительный интервал
ДР	– Диабетическая ретинопатия
ЗВИ	– Заболевания витреомакулярного интерфейса
MP	– Макулярный разрыв
HCO	<ul> <li>Неоваскуляризация сосудистой оболочки</li> </ul>
СД	– Сахарный диабет
CHC	– Сверточные нейронные сети
OAC	– Окклюзия артерий сетчатки
OBC	– Окклюзия вен сетчатки
ОКТ	– Оптическая когерентная томография
ПЭ	– Пигментный эпителий
ЦВС	– Центральная вена сетчатки
ЦСР	– Центральная серозная ретинопатия
ЭМ	– Эпиретинальная мембрана
AREDS	– Age-Related Eye Disease Study
ICDR	- International Classification of Diabetic Retinopathy
OpenCV	<ul> <li>Open Source Computer Vision Library</li> </ul>
OCTDL	- Optical Coherence Tomography Dataset for Image-Based Deep
	Learning Methods
ReLu	– Rectified Linear Unit
SVM	– Support Vector Machines

#### введение

Заболевания сетчатки второе причин занимают место среди инвалидности по зрению в России [1]. Такому заболеванию как возрастная макулярная дегенерация подвержено около 8 миллионов человек во всем мире согласно статистике Всемирной Организации Здравоохранения (ВОЗ) [2]. дегенерация **(ВМД)** Возрастная макулярная ЭТО хроническое котором прогрессирующее заболевание, при постепенно поражается центральная зона нервной ткани глаза — макулярная область сетчатки, или как ее еще называют, желтое пятно, отсутствие диагностики и дальнейшего лечения которого приводит к потере зрения [3].

Ha наибольшую сегодняшний день точность неинвазивного исследования срезов сетчатки и диагностики заболевания на ранних стадиях предлагает оптическая когерентная томография (ОКТ) – это высокоточный неинвазивный метод исследования структур глаза, основанный на технологии низкокогерентной интерферометрии. Однако, ручная оценка и интерпретация снимков ОКТ является чрезвычайно времязатратной процедурой, к тому же она подвержена субъективной оценке специалистов, что приводит к плохой воспроизводимости и снижает качество диагностики. Таким образом, проблема сегодняшний день остается актуальной на автоматизации диагностики офтальмологических заболеваний на снимках ОКТ. Однако, благодаря развитию и внедрению алгоритмов машинного и глубокого обучения, шанс не пропустить зарождение и самые ранние стадии заболевания, возрастает.

В данной работе объектом исследования являются снимки оптической когерентной томографии глаза с ранней и промежуточной формами возрастной макулярной дегенерации.

Предмет исследования – это разработка алгоритма нейронной сети для мультиклассовой сегментации «сухой» формы возрастной макулярной дегенерации.

Цель настоящей работы – разработать и протестировать модель нейронной сети, способную проводить мультиклассовую сегментацию ранней и промежуточной форм возрастной макулярной дегенерации на снимках оптической когерентной томографии глаза.

Для достижения поставленной цели необходимо выполнить следующие задачи:

1) провести предобработку снимков оптической когерентной томографии, включающую разметку, извлечение меток класса и изменение размера изображений;

2) оценить качество работы модели нейронной сети на основе полученных метрик;

3) сравнить полученные результаты с аналогичными работами в области применения глубокого обучения в диагностике заболеваний сетчатки глаза.

Научная новизна: в рамках работы впервые проведена мультиклассовая сегментация «сухой» формы ВМД на открытом наборе данных OCTDL с использованием сверточной глубокой нейронной сети UNet3+.

Практическая значимость работы заключается в том, что использование методов глубокого обучения для сегментации снимков ОКТ и классификации форм ВМД позволит диагностировать заболевание на ранней стадии и предотвратить его переход в более тяжелую форму. Также использование глубоких нейронных сетей сократит время обследования пациента врачом и снизит влияние субъективной оценки специалиста на постановку диагноза.

# 1 Исследование алгоритмов нейронных сетей для мультиклассовой сегментации ВМД на снимках ОКТ

### 1.1 Общая характеристика предметной области

С использованием платформы SciVal на основе базы данных Scopus в ходе анализа предметной области были рассмотрены библиометрические показатели за период 2018 – 2023 по следующей тематической области: Segmentation; Macular Edema; Speckle Noise. Статистика полученных показателей представлена на рисунке 1.



Рисунок 1 – Библиометрические показатели по тематической области: Eye; Segmentation; Macular Edema; Speckle Noise [4]

Scholary Output – количество научных публикаций организации или автора, которые проиндексированы в Scopus.

Field Weighted Citation Impact – взвешенный по области цитирования индекс – это отношение суммарного количества цитирований, фактически полученных работами всех авторов выбранной категории, к общему количеству цитирований, ожидаемому исходя из среднего показателя для соответствующей области.

International Collaboration – оценка международного сотрудничества, которая демонстрирует то, насколько часто авторы исследований из разных стран взаимодействуют для получения общего результата.

Views Count – количество просмотров научной статьи или другого материала, полученного в ходе исследования.

Citation Count – количество цитирований.

Торіс Prominence percentile – это процентиль значимости темы, который отражает значимость темы научного результата в сравнении с другими темами в данной области науки.

На рисунке 2 представлено облако слов, полученное по результатам поиска области исследования.

evice Retina Maculopathy Central Serous Retinopathy Semantic Segmentation Optical Coherence Tomography Device Cellular Neural Networks Spectral Domain Optical Coherence Tomography Disease Classification Eye Protection **Optical Data Processing** Subretinal Neovascularization Convolutional Neural Network Automatic Segmentation Multiscale omograpn d Diabetic Retinopathy Convolution Tomography Macular Edema Speckle Noise Deep Learning Speckle Reduction Retinopathy Diabetic Macular Edema Denoising Graph Search Image Denoising Macular Degeneration Glaucoma B Scan Generative Adversarial Networks Computer-aided Diagnosis Retina Disease Retinal Segmentation Coherent Light phthalmology Convolutional Network Image Noise Reduction Speckle Retina Image Retina Detachment Image Segmentation Multilayer Neural Network Subretinal Fluid Optical Coherence Tomography Angiography

Рисунок 2 – Облако слов, полученное по результатам поиска области исследования [4]

На основе полученных библиометрических показателей по тематической области можно сделать следующий вывод: динамика количества научных публикаций и международного сотрудничества положительная, а также взвешенный по области цитирования индекс выше среднего. Таким образом, выбранная тема исследования является актуальной, что объясняется активным развитием глубоких нейронных сетей и их применением для решения задач, связанных с медицинскими изображениями.

# 1.2 Анализ современных алгоритмов нейронных сетей для сегментации и классификации ВМД на снимках ОКТ

Для научного анализа были использованы исследовательские работы, в которых рассматривается применение алгоритмов нейронных сетей для решения задачи мультиклассовой сегментации возрастной макулярной дегенерации.

На наборе данных RETOUCH авторами было проведено исследование с целью определения наличия и сегментации трех классов скоплений жидкости в сетчатке при ВМД: внутриретинальной, субретинальной и отслоения пигментного эпителия (ПЭ) [5]. В качестве рабочего инструмента была использована модифицированная глубокая нейронная сеть U-Net, реализуемая в 2 этапа. Первый включает в себя добавление активации dropout и maxout на каждом слое, а также полносвязного слоя между слоями свертки и декодирования для повышения точности и предотвращения переобучения. В этом случае сеть принимает изображение ОКТ в качестве входных данных. Второй этап характеризуется удалением полносвязного слоя между слоями свертки и декодирования. Здесь в качестве входных данных сеть принимает и изображение-ОКТ, и соответствующее сегментированное изображение, сгенерированное на первом этапе. В ходе исследования были получены следующие результаты: Accuracy = 0.96, коэффициент Dice Coefficient = 0.90. Последний используется для сравнения попиксельного соответствия между прогнозируемой сегментацией и соответствующей ей достоверностью, чем он ближе к единице, тем точнее выполнена сегментация.

В исследовании К. Alsaih и др. использовалась выборка из 100 томов ОКТ-снимков из общедоступных наборов данных: RETOUCH (70 3Dизображений) и ОРТІМА (30 3D-изображений) для сегментации 3 типов жидкостей при ВМД [6]. В работе было проведено сравнение четырех алгоритмов нейронных сетей: FCNN, U-Net, Seg-net и PaDeeplabv3+. Последняя сеть, обладающая кодировщиком и декодером, позволяет преодолевать разницу между пикселями фона и заболеванием, и именно она

продемонстрировала наилучший результат для сегментации жидкости при ВМД, который составил Dice Coefficient = 0,84.

S. Sotoudeh-Paima и др. для классификации «сухой» и «влажной» форм ВМД была предложена многомасштабная сверточная нейронная сеть, основанная на структуре сети функциональных пирамид (FPN) [7]. Данная модель использует преимущества многомасштабных рецептивных полей, позволяя более точно выявлять патологии сетчатки разных масштабов на изображениях ОКТ. Этот метод обеспечивает сквозное обучение многомасштабной модели с помощью единой сверточной нейронной сети с использованием упрощенной конструкции и устраняет необходимость в обработке предварительной входных данных. Работа модели была протестирована на основе национального набора данных (NEH), собранного в больнице для этого исследования и состоящего из 12649 ОКТ изображений сетчатки и общедоступного набора данных UCSD, состоящего из 108312 ОКТ изображений. Наилучший продемонстрировала результат сверточная нейронная сеть FPN на базе VGG16 значение Accuracy = 0,92. Недостатком что он решает только данного метода является тот факт, задачу классификации.

Shengyong Diaoa и др. в своей работе предложили структуру глубокого обучения, решающую одновременно задачу сегментации и классификации [8]. Во-первых, предлагается дополнительная сверточная нейронная сеть, управляемая комплементарной маской для классификации снимков ОКТ с «сухой» и «влажной» формами ВМД и нормой. Полученная маска сегментации используется в дополнительной форме для извлечения признаков классификации и повышения ее эффективности. Во-вторых, предлагается UNet, управляемая картой активации классов, для сегментации двух форм ВМД, где карта активации классов из классификации объединяется с признаками на каждом уровне для решения задачи сегментации. В ходе тестирования, выполненного на общедоступном наборе данных UCSD, были получены следующие результаты: точность классификации (Classification

Accuracy) достигла 96,93% и коэффициент Dice для сегментации составил 77,51%.

В работе Х. Хі и др. была предложена сверточная нейронная сеть информационного внимания (IA-Net), основанная на базе сети UNet, для автоматической сегментации «влажной» формы ВМД на изображениях ОКТ [9]. Ее отличительными особенностями являются: блок повышения внимания и информативная карта внимания. Первый разрабатывается путем введения дополнительного ограничения внимания. Он обладает способностью заставить модель уделять повышенное внимание хориоидальной процессе обучения объектов, неоваскуляризации В карты улучшая различительную способность изученных функций малых объектов, что полезно для повышения точности сегментации заболевания на картах объектов низкого уровня и сохранения информации – на картах объектов высокого уровня. Это означает, что IA-Net может эффективно изучать больше отличительных признаков, что приводит к точной сегментации большего количества пикселей хориоидальной неоваскуляризации. Информативная карта внимания используется для получения точной классификации пикселей хориоидальной неоваскуляризации, которые трудно предсказать. В ходе исследования IA-Net продемонстрировала результат Dice Similarity Coefficient в 0,8862 и превзошла описанные в исследовании методы сегментации.

В работе R. Asgari и др. представлено альтернативное решение задачи многоклассовой сегментации друз, реализованное через сегментацию верхней и нижней анатомических границ, которые определяют друзы – внешнюю границу пигментного эпителия сетчатки и мембрану Бруха [10]. Используемая архитектура сети UNet разделена на k-декодеров, каждый из которых ориентирован на задачу бинарной сегментации. Одновременно, для описания общих характеристик между декодерами используется один кодировщик. Для захвата области между сегментируемыми слоями вместо того, чтобы иметь один декодер для каждого класса слоев, вводится третий для возможности кодировщику охарактеризовать появление не только друз, но И

непатологических областей. Стоит отметить, что входящие/исходящие соединения с декодером друз значительно повышают производительность, когда через них не пропускается градиентный поток. Это означает, что локальная информация, предоставленная соседними классами, может улучшить результаты. В работе использовался частный набор, состоящий из 560 ОКТ-снимков, который был вручную размечен. В связи с тем, что датасет включал в себя несколько сканирований одного и того же глаза пациента, метрики качества сначала рассчитывались для одного глаза пациента, а затем усреднялись по количеству исследуемых глаз. Были достигнуты следующие результаты: коэффициент Dice для декодера без градиентного потока составил 0,73, в то время как с градиентным потоком – 0,72.

Цель исследования R. Tennakoon и др. на наборе данных RETOUCH заключалась в определении наличия и сегментации трех классов скоплений жидкости в сетчатке при ВМД: внутриретинальной, субретинальной и отслоения пигментного эпителия [11]. Авторы статьи отмечают, что автоматизированная сегментация жидкости на изображениях ОКТ является сложной задачей в связи с высокой вариабельностью проявления патологии на изображениях, наличием шума и артефактов движения. Отличительной особенностью данного подхода является использование модификации глубокой нейронной сети U-Net со следующими особенностями:

– после каждого блока свертки (convolution)/деконволюции (deconvolution) добавляется слой пакетной нормализации (batch normalization layer), повышающий эффективность обучения;

– слой dropout используется в каждом пропускном соединении для предотвращения переобучения;

– последний уровень свертки U-Net использует только выходные данные последнего блока декодера.

В таблице 1 представлены основные характеристики исследований, описанных в анализе научных работ.

Ί	Габлина	1 – (	)сновные	характе	ристики	исспелова	аний
1	аолица	I (	Jenoblible	Aupunte	photnikh	песледов	ALLELE

N⁰	Статья	Алгоритм	Набор	Решаемая	Форма ВМД	Метрики
			данных	задача		качества
1	S. H. Kang,	Модифициров	RETOUCH	Сегментация,	«Влажная»	Accuracy
	H. S. Park, J.	анная UNet		классификация		= 0,96,
	Jang и др.	(UNet3+)				Dice
	[5]					Coefficient
						= 0,90
2	K. Alsaih,	PaDeeplabv3+	RETOUCH,	Сегментация	«Влажная»	Dice
	M.Z. Yusoff,		OPTIMA			Coefficient
	T.B. Tang и					= 0,84
	др. [6]					
3	S. Sotoudeh-	FPN-VGG16	NEH,	Классификация	«Сухая»,	Accuracy
	Paima, A.		UCSD		«Влажная»	= 0,92
	Jodeiri, F.					
	Hajizadeh и					
	др. [7]					
4	S. Diao, J.	Модифициров	UCSD	Сегментация,	«Сухая»,	Accuracy
	Su, C. Yang	анная UNet		классификация	«Влажная»	= 0,9693,
	и др. [8]	(UNet3+)				Dice
						Coefficient
_				~		=0,7751
5	X. Xi, X.	IA-Net	Частный	Сегментация	«Влажная»	Dice
	Meng, Z.		набор			Coefficient
	Qin и др. [9]		данных	~	~	=0,8862
6	R. Asgari, J.	Модифициров	Частный	Сегментация	«Сухая»	Dice
	I. Orlando,	анная UNet	набор			Coefficient
	F.G.		данных			= 0,73
	Schlanitz и					
-	др. [10]		DETOLICI	<u> </u>	D	D:
1	K.	Модифициров	RETOUCH	Сегментация	«Влажная»	Dice
	Tennakoon,	анная UNet				Coefficient
	A.K.	(UNet3+)				0,73
	Gostar, R.					
	Hoseinnezha					
	d и др. [11]					

Исходя из проведенного анализа научных работ, стоит отметить, что большая часть исследований посвящена диагностике поздней стадии ВМД, то есть «влажной» форме, однако наиболее ценной и актуальной на сегодняшний день является диагностика начальной и промежуточной стадий заболевания – «сухой» формы. Чем раньше оно выявлено, тем выше вероятность выздоровления. Также лишь небольшая часть исследований решает одновременно две задачи: сегментацию и классификацию, что снижает качество диагностики. Для решения задач, связанных с обработкой медицинских изображений, чаще всего используются сверточные нейронные сети. Операция свертки в них используется для построения карты признаков, в которой интенсивность каждого пикселя, вычисленная как сумма пикселя и всех его соседей, взвешивается на сверточных матрицах – ядрах свертки. Среди описанных в данной главе исследованиях наиболее высокие значения продемонстрировала модифицированная метрик модель сверточной нейронной сети UNet. Стоит отметить, что U-Net может показывать хорошее качество даже на небольших объемах данных (100-200 изображений). На основе анализа методов для данной работы было принято решение в качестве объекта исследования выбрать неэкссудативную форму ВМД, а в качестве инструмента – наиболее эффективный, но одновременно малозатратный по ресурсам алгоритм – глубокую сверточную нейронную сеть UNet3+.

# 1.3 Биомедицинские аспекты, связанные со строением сетчатки глаза, офтальмологическими заболеваниями и их диагностикой

#### 1.3.1 Строение сетчатки глаза

Ввиду нетривиального строения глаза офтальмологических проблем становится с каждым годом все больше. Также растущее количество заболеваний может свидетельствовать о том, что совершенствуются методы диагностики. На рисунке 3 представлена структура причин инвалидности по зрению в России за 1997 – 2016 год, из которой следует, что вторыми по распространенности являются заболевания сетчатки, причем за период в 20 лет их количество возросло практически в 3 раза.



Рисунок 3 – Структура причин инвалидности по зрению в России с 1997 по 2016 год [12]

Сетчатка, представляющая внутреннюю оболочку глазного яблока, – сенсорная часть зрительного анализатора, именно в ней под воздействием света происходят фотохимические превращения зрительных пигментов, фототрансдукция, изменение биоэлектрической активности нейронов и передача информации об окружающем мире в подкорковые и корковые зрительные центры. Сетчатка является одним из наиболее сложных в структурном и функциональном отношениях образованием глаза. На рисунке 4 представлено микроскопическое строение сетчатки.



Рисунок 4 – Микроскопическое строение сетчатки глаза [13]

Сетчатка состоит из 10 слоев клеток, которые истончаются при приближении к макуле:

1) пигментный эпителий и мембрана Бруха обеспечивают питанием фоторецепторы, доставляя соли, кислород и метаболиты из сосудистой оболочки, а также контролируют уровень жидкости между сетчаткой и сосудистой оболочкой для максимально плотного прилегания;

 фоторецепторный слой состоит из палочек, отвечающих за восприятие света и сумеречное зрение, и колбочек, от которых зависит острота зрения и восприятие цвета;

3) наружная пограничная мембрана разграничивает структуру палочек и колбочек, где внешний и внутренний сегменты фоторецепторов лежат в фоторецепторном слое, а тело клетки с ядром находится в наружном зернистом слое;

4) наружный зернистый (ядерный) слой образован из тел и ядер палочек и колбочек;

5) наружный сетчатый слой состоит из ретинальных капилляров и является первой ступенью превращения световых лучей в нервные импульсы;

6) внутренний зернистый (ядерный) слой содержит горизонтальные, биполярные, амакриновые, интерплексиформные клетки и клетки Мюллера, которые отвечают за обработку сигналов, полученных от фоторецепторов наружных слоев сетчатки;

7) внутренний сетчатый слой – последний уровень обработки информации внутри сетчатки, после которой она будет передана с помощью нервных клеток в мозг;

8) слой ганглиозных клеток генерируют нервные импульсы, которые поступают в другие структуры мозга;

9) слой волокон зрительного нерва обеспечивает возможность направлять созданные биоэлектрические импульсы в головной мозг;

10) внутренняя пограничная мембрана покрывает всю сетчатку и является границей между ее поверхностью и стекловидным телом [14].

#### 1.3.2 Возрастная макулярная дегенерация

Возрастная макулярная дегенерация является одной из основных причин слепоты (составляет 8,7 %) в развитых странах мира, и с каждым годом распространенность данного недуга возрастает [3]. Возрастная макулярная дегенерация — хроническое прогрессирующее заболевание, при котором постепенно поражается центральная зона нервной ткани глаза — макулярная область сетчатки, или как ее еще называют, желтое пятно. Заболеванию подвержены оба глаза, причем поражение парного глаза происходит в среднем в течение 5 лет после диагностирования на первом. К факторам риска развития ВМД относят возраст, курение, наследственную предрасположенность, соматические заболевания, такие как диабет, атеросклероз, артериальная гипертензия, европеоидную расу, вредные факторы окружающей среды (ультрафиолетовое излучение солнца) [15]. Менее негативное воздействие оказывает употребление пищи бедной антиоксидантами, избыточный вес, пол (обнаружение ВМД у мужчин происходит в два раза реже, чем у женщин) [16].

ВМД чаще всего развивается в пигментном эпителии, мембране Бруха, хориокапиллярном слое сетчатки, а с течением времени в патологический процесс вовлекаются фоторецепторы [15]. Одним из первичных признаков развития заболевания является формирование друз – образований, состоящих из вещества липофусцина, фрагментов жизнедеятельности старых клеток ПЭ и его остатков, а также липопротеидов В и Е [17]. Содержимое друз является стимулятором хронического воспаления, а фототоксичность липофусцина, при которой под действием света происходит генерация активных форм кислорода (АФК), негативно влияет на фоторецепторы. С течением развития заболевания, при переходе из ранней формы ВМД в промежуточную, одиночные друзы объединяются и становятся сливными. Рисунки 5 и 6 иллюстрируют примеры одиночных и сливных друз на снимках оптической когерентной томографии, соответственно.



Рисунок 5 – Одиночная друза на снимке ОКТ [18]



Рисунок 6 – Сливная друза на снимке ОКТ [18]

В настоящее время принято различать две формы ВМД: неэкссудативная или «сухая» и неоваскулярная или «влажная». Первая встречается наиболее часто и диагностируется в 80-90 % случаев.

Международная классификация ВМД по Age-Related Eye Disease Study (AREDS):

1) отсутствие ВМД (категория 1 AREDS) – контрольная группа в исследовании AREDS, для которой характерно отсутствие или малое количество друз диаметром < 63 мкм;

2) ранняя стадия ВМД (категория 2 AREDS, ранняя «сухая» форма ВМД) – характерны множественные мелкие друзы, небольшое количество друз среднего размера (диаметр 63-124 мкм), начальные изменения пигментного эпителия сетчатки;  промежуточная стадия ВМД (категория 3 AREDS, промежуточная «сухая» форма ВМД) – характерно большое количество друз среднего размера, хотя бы одна большая друза (диаметр ≥ 125 мкм), или географическая атрофия, не затрагивающая центральной ямки;

4) поздняя стадия ВМД (категория 4 AREDS) – при отсутствии других причин характеризуется одним или несколькими признаками, представленными ниже:

 – географической атрофией ПЭ и холиокапиллярного слоя в макулярной области;

- хориоидальной неоваскуляризацией;

- серозной или геморрагической отслойкой нейроэпителия или ПЭ;

- твердыми экссудатами;

– образованием дисковидного рубца, который является исходом неоваскулярной формы ВМД и представляет собой участок фиброза макулы [19].

Категории 1-3 по международной классификации AREDS, а также географическую атрофию пигментного эпителия из категории 4 AREDS относят к «сухой» форме возрастной макулярной дегенерации, остальные признаки – к «влажной». С развитием заболевания для «сухой» формы характерен переход из ранней формы ВМД в промежуточную, при котором одиночные друзы объединяются и становятся сливными. При дальнейшей кальцификации друз и заполнении их холестерином неэкссудативная форма ВМД перетекает в неоваскулярную. Для «влажной» формы характерными являются: атрофия прилегающих зон сетчатой оболочки, рост патологических сосудов (неоваскуляризация) ПЭ, процессы рубцевания, что приводит к потере значительного количества фоторецепторов (палочек и колбочек). Подобные проявления заболевания становится причиной заметного снижения остроты зрения.

#### 1.3.3 Гипертоническая ретинопатия

В исследовании И. В. Мунц и др. гипертоническая ретинопатия (ГР) занимает второе место по частоте встречаемости среди заболеваний сетчатки. ГР относится к сосудистым изменениям сетчатки, связанным с системной некомпенсируемой артериальной гипертензией [20]. В зависимости от степени тяжести ГР характеризуется такими проявлениями как сужение артерий сетчатки (I степень); артериовенозные трещины сетчатки (II степень); кровоизлияния в сетчатку, пятно перед глазом, твердый экссудат (III степень); отек диска зрительного нерва (ДЗН) (IV степень).

#### 1.3.4 Диабетическая ретинопатия

Диабетическая ретинопатия в работе И.В.Мунц и др. является третьим по распространенности среди заболеваний сетчатки [20]. ДР – это сосудистое заболевание сетчатки, являющееся основным осложнением сахарного диабета (СД) и встречающееся у 30-40% людей, страдающих этим недугом [21]. Его развитие происходит постепенно, начиная с повышенной проницаемости и непроходимости ретинальных сосудов до формирования новых сосудов и соединительной Согласно международной классификации ткани. диабетической ретинопатии (ICDR – International Classification of Diabetic ДР и Retinopathy) выделяют непролиферативную, пролиферативную диабетический макулярный отек (ДМО) [22]. Характерными признаками для первой являются микроаневризмы, внутриретинальные кровоизлияния, венозные интраретинальные аномалии, отсутствие И признаков пролиферативной ретинопатии. Пролиферативная форма характеризуется неоваскуляризацией, преретинальным кровоизлиянием. При разрастании новообразованных сосудов по радужке и структурам угла передней камеры сильно увеличивается вероятность развития неоваскулярной глаукомы. Диабетический макулярный отек подразумевает некоторое видимое

утолщение сетчатки или твердый экссудат в области глазного дна. Такие стадии как пролиферативная ДР и макулярный отек являются наиболее грозными и чаще приводят к полной потере зрения [23]. Основными факторами риска развития данного заболевания являются гипергликемия, артериальная гипертензия, дислипидемия, однако их сила и характер напрямую зависят от степени тяжести ДР.

# 1.3.5 Окклюзия центральной вены сетчатки и ветвей центральной вены сетчатки

Окклюзия центральной вены сетчатки (ЦВС) – это заболевание, сопровождающееся нарушением кровотока в центральной вене сетчатки и ее ветвях [24]. Природа его возникновения обусловлена сочетанием трех основных факторов: компрессией вены в области артериовенозного атеросклероза), перекреста (вследствие гипертонии, дегенеративными изменениями сосудистой стенки, нарушениями реологических свойств и фибринолитической активности крови (вследствие сахарного диабета). Коварность заболевания заключается в его бессимптомности. Среди возможных вариантов окклюзии ветви ЦВС выделяют две группы: окклюзию крупной ветви (верхне-височной, нижне-височной, верхне-носовой и нижненосовой) и окклюзию ветви, осуществляющей отток крови от макулярной зоны (нижняя и верхняя макулярная ветвь). Различают ишемический и неишемический типы нарушения кровообращения в ветвях ЦВС. В первом случае происходит ухудшение кровообращения в капиллярной сети сетчатки, приводящее к образованию «зон инфаркта сетчатки». И наоборот, при диагностировании неишемического тромбоза кровообращение в капиллярной сети сетчатки совсем или практически не нарушается. Такие проявления делают неишемический тип окклюзий более благоприятным, при этом частота неоваскулярных осложнений значительно снижается.

#### 1.3.6 Аппаратные методы диагностики заболеваний сетчатки глаза

Наибольший успех в лечении того или иного заболевания наблюдается, когда заболевание диагностировано на ранней стадии. Благодаря развитию технологий за последние годы это стало доступно большему количеству людей по всему миру. Так, наиболее востребованными аппаратными методами диагностики заболеваний сетчатки являются:

- визометрия - исследование остроты зрения;

 компьютерная периметрия – определение границы поля зрения и работы нервных рецепторов сетчатки во всех зонах поля зрения;

 прямая и обратная офтальмоскопия (осмотр глазного дна), а также осмотр сетчатки с помощью фундус-линзы;

томография – оптическая когерентная ЭТО высокоточный неинвазивный метод исследования структур глаза, основанный на технологии низкокогерентной интерферометрии. Спектральные оптические когерентные томографы работают по следующему принципу: световой импульс делится на 2 равные части, одна из которых отражается от фиксированного опорного плеча (зеркала), другая – от исследуемого объекта, далее сигналы суммируются, проинтерферировавший луч раскладывается на составные части спектра, одномоментно фиксирующиеся ССД-камерой, при этом фиксируется время этого процесса [25]. Из полученного массива данных, используя математическое преобразование Фурье, выделяются частотные составляющие, формируя А-сканы. Они, в свою очередь, позволяют выполнить реконструкцию срезов сетчатки, сосудистой оболочки И стекловидного тела в плоскости поперечного сечения, формируя В-сканы;

– оптическая когерентная томография с функцией ангиографии сосудов сетчатки глаза – это инновационный неинвазивный метод исследования сосудов центральной зоны глазного дна – сетчатки и хориоидеи, позволяющий, кроме того, оценивать состояние сосудистой сети диска зрительного нерва [26].

2 Архитектура нейронной сети UNet3+ для мультиклассовой сегментации ВМД на снимках ОКТ

# 2.1 Выбор подхода для решения задачи мультиклассовой сегментации

Машинное обучение – класс методов искусственного интеллекта, применяемых для анализа сложных данных и нахождения паттернов и взаимозависимостей без их явного программирования [27]. Другим способом реализации искусственного интеллекта является искусственная нейронная сеть. Она представлена слоями из связанных между собой узлов. Узел содержит сумму весов, которая передается на функцию активации по выходной связи, а затем на следующий узел. В процессе обучения нейронной сети веса признаков динамически изменяются. Благодаря эффективной работе со сложными и многомерными базами данных, возрастающей доступности графических наборов данных И производительности процессоров использование нейронных сетей на основе глубокого обучения становится все более востребованным.

Выбор подходящей модели для решения задачи мультиклассовой сегментации, реализуемой в данной работе, основан на возможности решения одновременно 2 задач – классификации и сегментации, соответственно. Задача классификации может быть выполнена с помощью методов классического обучения, таких машинного как: Логистическая регрессия (Logistic Regression), Дерево решений (Decision Tree), Случайный лес (Random Forest), Метод опорных векторов (SVM – Support Vector Machines), Наивный Байесовский классификатор (Naive Bayes classifier), метод К-ближайших соседей (K-Nearest Neighbors), и ансамблевых методов – AdaBoost, XGBoost. Однако, этот выбор сильно сужается, если говорить о задаче сегментации. Среди классических моделей машинного обучения с ней способны справиться только: Случайный лес, Метод опорных векторов и Логистическая регрессия.

В работе Р. Serrano-Aguilar Р. и др. был разработали метод автоматизированной диагностики неоваскулярной ВМД на основе выборки обучающих данных, состоящей из 260 изображений ОКТ (136 – ВМД, 124 – норма) [28].

В качестве используемого алгоритма выступали Деревья решений, которые оказались полезны для определения характеристик нормальных и аномальных изображений с использованием векторов признаков, полученных из описательной статистики обнаруженных структур. Результатом исследования являются следующие показатели: Чувствительность (Sensitivity) = 96 % и Специфичность (Specificity) = 92 %.

R. J. Zawadzki и др. в своем исследовании используют метод опорных векторов для выполнения полуавтоматической сегментации слоев и структур сетчатки глаза для последующего анализа, включая сравнение измеренных толщин слоев [29]. Данная модель машинного обучения используется для прогнозирования результатов на основе заданного набора входных данных. Преимуществами использования SVM является возможность обработки спекл-шума, который мешает данным, полученным с помощью ОКТ, моделируя их как нормальное распределение. Данный метод допускает неверную классификацию пользователем и физиологические различия между пациентами и заболеваниями и легко адаптируется к различным свойствам данных, составляющих слой сетчатки. Описываемый подход был эффективен при оценке трехмерных структур сетчатки как у здоровых, так и у больных пациентов, что позволило облегчить диагностику и мониторинг лечения заболеваний сетчатки. Недостатком данного исследования является тот факт, что демонстрация результатов работы выбранного алгоритма реализована на конкретных снимках, но в исследовании отсутствуют метрики, позволяющие объективно оценивать эффективность работы.

S. J. Chiu и др. был предложен полностью автоматический метод классификации на основе ядерной регрессии для идентификации заполненных жидкостью областей и семи слоев сетчатки на изображениях оптической

когерентной томографии в спектральной области глаз с диабетическим макулярным отеком [30]. Алгоритм был протестирован на 110 В-снимках, полученных от десяти пациентов с тяжелой патологией ДМО, и вычислен общий средний коэффициент Dice, равный 0,78. Этот метод сегментации – впервые полностью автоматизированный и семислойный из примененных ранее к реальным изображениям, содержащим тяжелые формы ДМО.

Глубокое обучение – метод, который использует многослойные нейронные сети для автоматического извлечения высокоуровневых признаков из сложных данных. Методы, основанные на глубоком обучении, требуют больших объемов данных, но они более продуктивны. Часто задача сегментации состоит из нескольких компонентов: снижения шума на изображениях ОКТ, а также сегментации слоев сетчатки. Еще одним преимуществом методов глубокого обучения является то, что многие системы, основанные на глубоком обучении, имеют сквозную системную архитектуру, в которой для решения основных компонентов задачи сегментации используется одна нейронная сеть.

Обобщая вышесказанное, можно сделать вывод, что классические алгоритмы машинного обучения обеспечивают гибкость и интерпретируемость, а алгоритмы глубокого обучения обеспечивают мощную масштабируемость и возможности иерархического изучения признаков, которые могут автоматически извлекать морфологические признаки из необработанных данных изображения. Таким образом, алгоритмы глубокого обучения на основе нейронных сетей демонстрируют лучшие результаты и больше подходят для решения задачи мультиклассовой сегментации объектов на медицинских изображениях, чем классическое машинное обучение.

#### 2.2 Сверточные нейронные сети

Сверточные нейронные сети (СНС) (Convolutional Neural Networks) – определенный вид нейронных сетей, специализирующийся на обработке

изображений, видео, звука, который имеет особую архитектуру, позволяющую обрабатывать большие объемы данных [31]. Отличительной особенностью СНС является их способность автоматически извлекать признаки из изображений. Благодаря обучению нейронной сети на огромном количестве различного качества изображений, имеющих свой определенный класс, это становится возможным. На рисунке 7 изображена упрощенная структура сверточной нейронной сети.



Рисунок 7 – Упрощенная структура сверточной нейронной сети [32]

Сверточная нейронная сеть начинается входным слоем (Input layer), на который поступают входные данные, например, снимок. Изображение является массивом данных, в котором каждый пиксель содержит значения, описывающие его цвет и интенсивность.

Основным элементом сверточной нейронной сети является сверточный слой (Convolutional layer), который применяет к входным данным операцию свертки вместо обычного умножения матриц по крайней мере на одном из своих уровней. Свертка (Cnvolution) – математическая операция, суть которой заключается в перемещении фильтра, или как его еще называют, ядра свертки (на рисунке 7 выделен белым квадратом), по всей области входных данных и вычислении скалярного произведения между фильтром и соответствующей областью входных данных. В результате применения операции свертки формируется новая матрица, которая содержит информацию о тех признаках изображения, которые были выделены фильтром. Предназначение операции

свертки заключается в выделении признаков изображения перед классификацией.

После слоя свертки чаще всего следует слой уменьшения размерности или субдискретизирующий слой (Pooling layer/Subsampling layer) [33]. Он применяется для уменьшения размерности карты признаков. Уменьшение размерности бывает двух видов: AveragePooling и MaxPooling. В первом случае рассчитывается среднее значение по выбранному окну, что является полезным для извлечения более общей информации. В случае применения MaxPooling выбирается максимальное значение из каждой области, при этом наиболее выделяются значимые признаки. Операция уменьшения размерности операцией свертки следующую схожа С И имеет последовательность: выбирается скользящее окно, которое двигается по карте признаков, в нем выбирается максимальное или среднее значение, и формируется карта признаков меньшего размера. На рисунке 8 представлен пример применения операции Pooling.



Рисунок 8 – Пример применения операции Pooling [33]

Слой выпрямления (Flatten layer) преобразует многомерные данные в одномерный вектор и подготавливает их к классификации. Это позволяет обеспечивать связь между сверточными слоями и полносвязными и повышать эффективность параметров [34].

Полносвязный слой (Fully Connected/Dense Layer) необходим для принятия моделью решения о принадлежности изображения к определенному классу. Каждый нейрон полносвязного слоя связан со всеми выходами предыдущего слоя Flatten, что позволяет комбинировать все высокоуровневые признаки для решения определенных задач, таких как классификация.

Завершает строение СНС выходной слой (Output layer), который обычно использует функцию активации для вычисления вероятности принадлежности входного изображения к каждому из возможных классов. В конечном счете получается распределение вероятностей по всем классам, это предоставляет возможность сделать окончательное предсказание.

Функция активации определяет выходное значение нейрона в зависимости от результата взвешенной суммы входных сигналов и порогового значения. Ее задача заключается в добавлении нелинейности в нейронную сеть, что позволяет модели обучаться и адаптироваться к новым данным [35]. Немаловажным является тот факт, что использование функции активации позволяет избегать проблему затухающего или взрывающегося градиента, что является ключевым в успешном обучении глубоких нейронных сетей.

В качестве функции активации в сети UNet используется ReLu (Rectified Linear Unit), вычисляемая по формуле (1):

$$f(x) = \max(0, x), \tag{1}$$

где f(x) – это значение функции активации ReLu,

х – входное значение функции.

Функция ReLu возвращает входное значение x, если x > 0, возвращает 0, если x ≤ 0. График функции активации ReLu представлен на рисунке 9.



Рисунок 9 – График функции активации ReLu [35]

Таким образом, функция активации ReLu получила широкое распространение и хорошо зарекомендовала себя благодаря следующим преимуществам:

 – устранение проблемы затухающего градиента, который не сходится к нулю при больших положительных значениях, что способствует ускорению обучения глубоких нейронных сетей;

 увеличение разреженности активаций в нейронной сети реализуется путем обнуления отрицательных входов, что способствует улучшению эффективности и уменьшению переобучения;

 эффективность вычислений требует меньше вычислительных ресурсов, поскольку включает в себя простые операции сравнения и присвоения;

 хорошие практические результаты при использовании в глубоких нейронных сетях.

#### 2.3 Алгоритм работы нейронной сети UNet и её модификации

UNet – полностью сверточная нейронная сеть на основе глубокого обучения с архитектурой кодировщика-декодера, предложенная О. Ronneberger и др. в 2015 году для решения задачи сегментация нейронных

структур в наборах изображений электронной микроскопии [36]. На рисунке 10 представлена архитектура нейронной сети UNet, которая напоминает букву «U» благодаря симметричности пути сжатия (левой части) и пути восстановления (правой части) изображения.



Рисунок 10 – Архитектура нейронной сети UNet [36]

Основными составляющими сети UNet являются: энкодер, декодер и связующий слой. Задача энкодера заключается в сжатии и уменьшении размерности входного изображения благодаря использованию связующего слоя, состоящего из слоев свертки, и извлечении признаков из него. Каждый слой свертки включает несколько фильтров, сканирующих изображение и выделяющих определенные признаки. После каждого слоя свертки применяется функция активации. Сеть имеет всего 23 сверточных слоя.

Задача декодера заключается в восстановлении изображения из сжатого представления. Транспонированные слои свертки, которыми образован декодер, увеличивают размерность изображения и восстанавливают его форму. Каждый из них состоит из нескольких фильтров, которые сканируют изображение и выделяют определенные признаки. В декодере так же как и в энкодере после каждого транспонированного сверточного слоя применяется

функция активации. Информация о признаках в процессе декодирования передается от энкодера декодеру с помощью связей между слоями, что способствует повышению качества сегментации. Слой с функцией активации Softmax преобразует выходные данные в вероятности для каждого пикселя изображения и используется в конце декодера. Вероятности предназначены для сегментации изображения на классы.

Отличительной особенностью UNet является наличие skip-connections – пропускающих соединений, которые позволяют передавать больше семантической информации низкого уровня о признаках из энкодера в декодер, пропуская некоторые слои, что делает результаты более точными. Пропускающие соединения играют важную роль в сохранении информации о мелких деталях изображения, которые могут быть утеряны при уменьшении размерности изображения в энкодере. Skip-connections отмечены на рисунке 10 серыми стрелками.

Сеть UNet имеет некоторое количество модификаций, которые отличаются характеристиками и количеством слоев. Так, сеть Unet++ устраняет недостатки предшествующей ей сети Unet [37]. В архитектуре Unet++ реализован глубокий контроль, также в ней части энкодера и декодера соединены с помощью вложенных, плотных пропускных соединений и блоков. Модифицированные пропуска агрегирующих пути через агрегирующие блоки направлены на сокращение семантического разрыва между картами объектов подсетей энкодера и декодера перед слиянием. Количество агрегирующих блоков пропорционально квадрату глубины сети. Модель может более эффективно захватывать мелкие детали объектов переднего плана, когда карты объектов с высоким разрешением из сети энкодера постепенно обогащаются перед объединением с соответствующими семантически богатыми картами объектов из сети декодера. Глубокий контроль позволяет модели работать в двух режимах: 1) точный режим, в котором выходные данные всех ветвей сегментации усредняются; 2) быстрый режим, в котором итоговая карта сегментации выбирается только из одной из

ветвей сегментации, выбор которой определяет степень сокращения модели и прирост скорости. Благодаря изменению структуры архитектура сети Unet++ эффективна и обеспечивает значительный прирост производительности.

На смену сети Unet++ пришла усовершенствованная сеть Unet3+, которая использует преимущества полномасштабных пропускных соединений (full-scale skip connections) и глубокий контроль [38]. Первые включают в себя детали низкого уровня с семантикой высокого уровня из карт объектов в разных масштабах, но с меньшим количеством параметров. Глубокий контроль изучает иерархические представления на основе полномасштабных агрегированных карт объектов, которые оптимизирует гибридную функцию потери данных для улучшения привязки к органу. Данная сеть включает модуль, основанный на классификации, для уменьшения чрезмерной сегментации изображений, не относящихся ни к одному органу, путем совместного обучения с классификацией на уровне изображений.

Сравнивая между собой 3 модификации сети Unet, можно отметить: Unet++ является более точной, чем Unet за счет содержания большей информации о функциях, но она состоит из большего количества параметров, что снижает скорость обучения модели. Unet3+ же сохраняет преимущества Unet++, при этом из ее структуры удаляются промежуточные узлы – блоки агрегации, что снижает количество параметров модели и делает ее более лаконичной. Полномасштабные пропускные соединения в Unet3+ вносят изменения во взаимосвязи между энкодером и декодером, а также внутренними соединениями между подсетями декодера, позволяя каждому уровню декодера объединять мелкомасштабные и одномасштабные карты объектов из энкодера и крупномасштабные карты объектов из декодера, что позволяет полномасштабно захватывать мелкозернистую семантику и крупнозернистую семантику. Наличие модуля классификации в Unet3+ позволяет избежать ложных срабатываний, вызванных зашумленными данными и чрезмерной сегментацией объектов на медицинских изображениях.

Таким образом, предлагаемая архитектура Unet3+ является малозатратной и высокопроизводительной и превосходит все предыдущие модификации сети Unet, сегментируя объекты разных масштабов и создавая четкие границы. Внешний вид архитектур нейронных сетей Unet, Unet++ и Unet3+ представлен на рисунке 11.



Рисунок 11 – Архитектуры нейронных сетей Unet, Unet++ и Unet3+ [36]

#### 2.4 Метрики оценки качества работы модели нейронной сети

Для оценки качества работы модели нейронной сети используются метрики качества. Так, качество сегментации можно оценить с помощью Dice Coefficient и IoU (Inception Over Union), классификации – с помощью Accuracy.

Dice Coefficient используется для сравнения попиксельного соответствия между прогнозируемой сегментацией и соответствующей ей достоверностью, чем он ближе к единице, тем точнее выполнена сегментация [39]. Dice Coefficient является аналогом F1-score и вычисляется по формуле (2):

Dice Coefficient = 
$$2 \times TP/(2 \times TP + FP + FN)$$
, (2)

где ТР (True Positive) – количество пикселей, корректно отнесенных к классу объекта,

FP (False Positive) – количество пикселей, ошибочно отнесенных к классу объекта,

FN (False Negative) – количество пикселей, ошибочно отнесенных к классу фона.

Значения Dice Coefficient варьируются от 0 до 1, чем значение ближе к единице, тем точнее выполнена сегментация.

Другой метрикой сегментации является Jaccard metric или, как ее еще называют, IoU (Intersection Over Union). Она используется для оценки сходства между двумя наборами данных и вычисляется по формуле (3):

$$IoU = TP/(TP + FP + FN),$$
 (3)

где ТР (True Positive) – количество пикселей, корректно отнесенных к классу объекта,

FP (False Positive) – количество пикселей, ошибочно отнесенных к классу объекта,

FN (False Negative) – количество пикселей, ошибочно отнесенных к классу фона.

Аналогично Dice Coefficient значения IoU варьируются от 0 до 1, причем единица означает полное совпадение наборов данных, а 0 – полное несоответствие.

Ассигасу – метрика классификации, отражающая долю правильных ответов алгоритма. Она вычисляется по формуле (4):

$$Accuracy = (TP + TN)/(TP + TN + FP + FN),$$
(4)

где ТР (True Positive) – количество пикселей, корректно отнесенных к классу объекта,

TN (True Negative) – количество пикселей, корректно отнесенных к классу фона,

FP (False Positive) – количество пикселей, ошибочно отнесенных к классу объекта,

FN (False Negative) – количество пикселей, ошибочно отнесенных к классу фона.

Данная метрика не является достаточно информативной для классификации изображений ввиду значительной разницы размеров фона и исследуемого объекта, особенно для медицинских изображений.

#### 2.5 Выбор инструментов для разработки

2.5.1 Выбор программного стека для обучения нейронной сети UNet3+

В качестве языка программирования был выбран высокоуровневый язык общего назначения Python. Он широко используется для анализа данных, машинного обучения и алгоритмов нейронных сетей. Наличие у него большого количества библиотек и фреймворков упрощает работу с данными.

Основными преимуществами языка Python являются:

 высокая популярность языка, в результате чего имеет множество библиотек и расширений, а также, высокую поддержку сообщества разработчиков;

 имеет понятный и простой синтаксис, что отлично сказывается на низком пороге входа, улучшает читаемость и снижает затраты на обслуживание кода;

– наличие встроенных библиотек, таких как Numpy, Tensorflow, OpenCV делают Python отличном инструментом для работы с изображениями [40].

Для обучения нейронной сети определенным задачам необходимо выбрать подходящие библиотеки. Keras – открытая библиотека для машинного и глубокого обучения, которая является надстройкой над основной библиотекой TensorFlow [41]. Задача последней – выполнение низкоуровневых сложных математических вычислений и преобразований.

Задача Keras – управление моделями, по которым распространяются вычисления.

Преимущества Keras:

– гибкое взаимодействие с другими библиотеками;

 – удобство использования – для решения наиболее часто возникающих задач необходимо выполнение минимального количества действий;

 модульное строение позволяет разработчику собирать из готовых элементов (слоев нейронных сетей, функций, схем) модель, что сильно сокращает время построения требуемого алгоритма;

- возможность написания собственного модуля;

- высокая скорость работы.

При работе с нейронными сетями каждый разработчик сталкивается с большими вычислениями, поддержку которых может оказать библиотека NumPy (Numerical Python), созданная для работы с большими многомерными массивами и матрицами [42].

Основные достоинства NumPy:

 высокая скорость вычислений достигается благодаря эффективным функциям и методам для работы с массивами, оптимизированным для быстрого выполнения;

 – удобство использования заключается в простых и понятных функциях для работы с массивами и матрицами, что облегчает понимание и реализацию алгоритмов;

– возможность интеграции с другими библиотеками.

Для решения задач, связанных с компьютерным зрением и обработкой изображений, отлично подходит открытая библиотека OpenCV (Open Source Computer Vision Library) [43]. К аргументам в пользу использования данной библиотеки можно отнести:

– бесплатный доступ;

 открытый исходный код предоставляет большую гибкость в работе и позволяет самостоятельно узнать, как реализована определенная функция;

- большое количество алгоритмов;

высокая скорость работы позволяет в короткие сроки обработать изображение;

- активное сообщество пользователей.

Одной из наиболее важных библиотек для визуализации данных является комплексная библиотека Matplotlib. Ее предназначение заключается в создании статических, анимированных и интерактивных визуализаций на языке Python [44].

К основным преимуществам библиотеки Matplotlib относятся:

– универсальность – возможность построения большого разнообразия
 графиков, таких как линейные, точечные графики, гистограммы, круговые
 диаграммы;

 настройка отображения графика – выбор необходимого стиля, цвета линий, маркеров, меток;

– интеграция с NumPy упрощает построение массивов данных напрямую;

– кроссплатформенность – возможность работы на различных операционных системах, таки как Windows, macOS и Linux [45].

Таким образом, в данной работе используется язык программирования Python и связанные с ним библиотеки: Keras, Numpy, OpenCV и Matplotlib.

#### 2.5.2 Выбор обучающей платформы

Среди обучающих платформ наибольшей популярностью пользуются Google Colab и Kaggle. Каждая из них имеет как свои плюсы, так и минусы.

Google Colab (Colaboratory) – это облачная среда, разработанная компанией Google, которая предоставляет пользователю возможность работать в браузере с кодом на языке Python через Jupyter Notebook, не устанавливая на компьютер дополнительных программ [46-47].

Преимущества Google Colab:

- бесплатный сервис;

 – система является облачной, что позволяет вести командную работу над одним проектом;

– возможность подключения к Google Диск – бесплатной облачной платформе для хранения файлов и безопасного предоставления доступа к ним;

– есть доступ к графическим процессорам GPU и TPU, чьи мощности активно используются для работы с глубокими нейронными сетями.

Недостатки Google Colab:

 дополнительные вычислительные мощности можно получить с платной подпиской Premium;

ограниченное время выполнения каждой сессии (12 часов) и
 ограниченный объем доступной оперативной памяти (около 12 ГБ);

– возможность работы только с одним языком программирования – Python.

Kaggle – это обучающая платформа, позволяющая пользователям находить или публиковать наборы данных, разрабатывать и обучать модели машинного обучения или нейронные сети, взаимодействовать с другими ML-специалистами и соревноваться с ними в решении задач [48-49].

Среди популярных соревнований на Kaggle стоит выделить задачи по обнаружению различных болезней [50]. Kaggle предоставляет возможности для работы с изображениями, поскольку на платформе проводятся соревнования по классификации, сегментации и другим задачам обработки изображений. Участники могут применять разнообразные архитектуры нейронных сетей, включая U-Net, для решения этих задач. Также на Kaggle доступны многочисленные обучающие материалы и решения задач, способствующие улучшению навыков работы с изображениями и машинным обучением в целом.

Преимущества Kaggle:

- удобный интерфейс;

– возможность использования графического процессора GPU для обработки больших объемов данных и ускорения обучения нейронных сетей;

 наличие обучающих курсов на платформе позволяет новичку быстро освоить основы машинного обучения и нейронных сетей;

 возможность участия в соревнованиях по решению актуальных проблем на реальных наборах данных.

Недостатки Kaggle:

– ограниченное время бесплатного использования GPU – 30 часов в неделю.

На практике было проведено сравнение скорости обучения глубокой нейронной сети с использованием двух обучающих платформ, в ходе которого определилось, что процесс обучения занимает в 3 раза меньше времени на платформе Kaggle. В связи с этим было принято решение для обучения нейронной сети использовать платформу Kaggle.

#### 2.6 Набор данных для обучения

### 2.6.1 Анализ обучающих наборов данных

Для разработки и оценки эффективности методов машинного обучения в области обработки медицинских изображений используют нормализованные базы данных, которые делятся на наборы открытого и ограниченного доступа [27]. К наборам открытого доступа относятся: RETOUCH, OPTIMA, UCSD, OCTID, OCTDL.

RETOUCH предназначен Набор данных для обнаружения И сегментирования трех типов жидкостей сетчатки: внутриретинальной жидкости, субретинальной жидкости и отслоения пигментного эпителия, что позволяет диагностировать такие заболевания как «влажная» форма ВМД, окклюзия вен сетчатки (OBC) [6]. Он состоит из 70 3D-изображений ОКТ, полученных с помощью оптических когерентных томографов: Cirrus, Triton, Spectralis вручную В-сканы были И размеченных специалистами.

аннотированы в Венском медицинском университете и Медицинском центре Университета Радбауд. Набор данных RETOUCH активно используется во многих исследованиях, связанных с сегментацией и классификацией жидкости сетчатки.

ОРТІМА – набор данных, полученный в результате исследования по сегментации интраретиальных кист, являющихся проявлением «влажной» формы ВМД [51]. Он состоит из 30 маркированных двумя экспертамиоценщиками из Христианской допплеровской лаборатории Венского медицинского университета 3D-изображений ОКТ, созданных с помощью оптических когерентных томографов: Cirrus, Topcon, Spectralis и Nidek. Набор данных был разделен на обучающую и тестовую выборки с помощью сканирований макулы. Недостатком описываемого набора данных является точная локализация областей сегментации интраретинальных кист, содержащихся в изображениях, полученных с разных устройств.

Набор данных ОКТ сетчатки UCSD содержит 108312 ОКТ-изображений от 4686 пациентов [52]. Изображения были получены с помощью ОКТ Spectralis в Heidelberg Engineering, Германия. UCSD состоит из четырех категорий: неоваскуляризация сосудистой оболочки (НСО), диабетический макулярный отек, «сухая» форма ВМД (друзы) и норма. Набор данных был собран из ретроспективных групп взрослых пациентов различными Глазной институт Шайли Калифорнийского учреждениями, включая университета в Сан-Диего, Калифорнийский фонд исследований сетчатки, Медицинский центр Ophthalmology Associates, Шанхайскую первую народную больницу и Пекинский офтальмологический центр Тунжэнь в период с 1 июля 2013 года по 1 марта 2017 года. UCSD обладает большим размером выборки, но значительным дисбалансом между категориями заболеваний.

Открытый набор данных OCTID, составленный в офтальмологической больнице Шанкара Нетралайя, Ченнаи, Индия, содержит более 500 3D-OKT-изображений с высоким разрешением, разделенных на категории по

различным патологическим состояниям [53]. Набор данных включает макулярный разрыв (МР), ВМД, центральную серозную ретинопатию (ЦСР) и диабетическую ретинопатию, а также норму. В ОСТІО были включены 25 обычных изображений вместе с соответствующими точными границами, которые были определены экспертом. Описываемый набор данных служит ценным ресурсом для ранней диагностики и мониторинга заболеваний сетчатки.

Набор данных OCTDL (Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods) содержит 2064 растровых OKT-изображения в формате .jpg с осевым и поперечным разрешением, равным 5 мкм и 15 мкм соответственно, полученных от 821 пациента и размеченных в соответствии с группой заболеваний и патологией сетчатки [54]. Изображения были получены с использованием устройства Optovue Avanti RTVue XR с применением протоколов растрового сканирования с динамической длиной сканирования и разрешением изображения. Каждое В-сканирование сетчатки было получено путем центрирования по ямке, а затем интерпретировано и каталогизировано опытным специалистом по сетчатке. В таблице 2 представлены категории заболевания и соответствующее им количество снимков.

N⁰	Заболевание	Количество ОКТ-изображений
1	Возрастная макулярная дегенерация	1231
2	Диабетический макулярный отек	147
3	Эпиретинальная мембрана (ЭМ)	155
4	Норма	332
5	Окклюзия артерии сетчатки (ОАС)	22
6	Окклюзия вен сетчатки	101
7	Заболевания витреомакулярного интерфейса (ЗВИ)	76
	Итого:	2064

Таблица 2 – Состав набора данных ОСТDL

Снимки были предоставлены специалистами из офтальмологической клиники «Профессорская плюс», г. Екатеринбург. Набор данных ОСТDL появился в открытом доступе и стал общедоступным в декабре 2023 года и обновлен в марте 2024 года [18]. Он был выпущен для исследований и разработки алгоритмов и предлагает полностью размеченные изображения для автоматической обработки и раннего выявления офтальмологических заболеваний.

Помимо наборов открытого доступа существует большое количество наборов данных с ограниченным доступом.

Сбалансированный набор данных для диагностики макулярной жидкости в работе T. Schlegl и др. [55] состоит из 1200 3D-изображений ОКТ пациентов с «влажной» формой ВМД, ДМО и окклюзией вен сетчатки, сделанных с помощью ОКТ-устройств Cirrus и Spectralis и представленных в базе данных Vienna Reading Center. Изображения были сегментированы двумя опытными специалистами по сетчатке.

F. G. Venhuizen и др. из Европейской Генетической базы данных был собран набор данных из 221 3D-изображения ОКТ, состоящих из 6158 Всканирований, от 151 пациента [56]. Необходимость его создания заключалась в сегментации внутриретинальной жидкости при ВМД. Снимки были сделаны при помощи устройства Spectralis. Стоит отметить, что алгоритм нейронной сети, разработанный авторами исследования, повысил эффективность сегментации объектов при использовании этого набора данных за счет внедрения в априорной информации об анатомии сетчатки.

Набор данных Bao D., Cheng X., Zhu W. и др. состоит из 240 Всканирований для сегментации отслойки пигментного эпителия, которая является крайним проявлением «влажной» формы ВМД [57]. Сложность сегментации данной стадии заболевания заключается в серьезной неравномерности размеров отслойки пигментного эпителия.

В таблице 3 представлены основные характеристики описанных наборов данных.

Набор данных	Объем набора	Используемый	Заболевание	Доступность
	данных			
		когерентный		
		томограф		
RETOUCH [6]	70 3D-	Cirrus, Triton,	«Влажная» ВМД,	Открытый
	изображений ОКТ	Spectralis	ОКС	
OPTIMA [51]	30 3D-	Cirrus, Topcon,	«Влажная» ВМД	Открытый
	изображений ОКТ	Spectralis и		
		Nidek		
UCSD [52]	108 312	Spectralis	НСО, ДМО,	Открытый
	изображений ОКТ		«сухая» ВМД,	
			норма	
OCTID [53]	500 3D-	_	МД, ВМД, ЦСР,	Открытый
	изображений ОКТ		ДР, норма	
OCTDL [54]	2064 изображения	Optovue Avanti	ВМД, ДМО, ЭМ,	Открытый
	OKT	RTVue XR	норма, ОАС,	_
			ОВС, ЗВИ	
T. Schlegl и	1200 3D-	Cirrus,	«Влажная» ВМД,	Закрытый
др. [55]	изображение ОКТ	Spectralis	ДМО, ОВС	
F. G.	221 3D-	Spectralis	«Влажная» ВМД	Закрытый
Venhuizen и	изображение ОКТ	_		
др. [56]	_			
D. Вао и др.	240 изображений	_	«Влажная» ВМД	Закрытый
[57]	OKT			

Таблица 3 – Характеристики наборов данных ОКТ-изображений

На основе информации о наборах данных, представленной в научных работах, а также характеристик, отраженных в таблице 3, в настоящей работе был выбран набор данных ОСТDL для обучения алгоритма UNet3+, при этом были учтены все его преимущества и недостатки. Сильными сторонами ОСТDL являются:

– доступность – набор данных является открытым;

– большое количество обучающих данных – 2064 снимка ОКТ;

 – 7 категорий заболеваний открывают дополнительные возможности для получения более точных результатов сегментации.

#### 2.6.2 Разметка обучающих наборов данных

Для качественной работы с архитектурой нейронной сети было необходимо разметить снимки ОКТ по двум стадиям ВМД, начальной и

промежуточной, то есть по двум классам – друзы и сливные друзы. Для решения данной задачи был выбран инструмент Labelme [58]. Labelme – это инструмент для создания графических аннотаций к изображениям, написанный на языке Python. Аннотации сохраняются в виде файла формата .json. Разметка проводилась с использованием операционной системы Windows 10.

Интерфейс Labelme и исходный вид изображения ОКТ в формате .jpeg до разметки представлены на рисунке 12.



Рисунок 12 – Интерфейс Labelme и снимок ОКТ до разметки

С помощью кнопки «Create Polygons» была обведена друза, верхняя часть которой находится под пигментным эпителием, а нижняя касается мембраны Бруха. Далее ей был присвоен один из классов (drusen – друзы, draining\_drusen - сливные друзы). Реализация представлена на рисунке 13.



Рисунок 13 – Присвоение класса сегментированному объекту

На рисунке 14 представлен окончательный вариант разметки всех объектов на изображении ОКТ. Красным цветом размечены одиночные друзы, зеленым – сливные друзы.



Рисунок 14 – Разметка друз на снимке ОКТ

Результатом разметки с помощью инструмента Labelme являются 160 размеченных ОКТ-изображений в формате .json, содержащих метку класса. Средний размер изображений составляет 800 х 300 пикселей, изображения в оттенках серого.

### 3 Обучение модели глубокой сверточной нейронной сети UNet3+ для решения задачи мультиклассовой сегментации ранней и промежуточной форм ВМД

Посредством языка Python и библиотеки OpenCV были получены маски, при этом координаты полигонов 160-ти размеченных изображений были преобразованы из формата .json в изображения формата .png с сохранением размеченных классов. Каждый класс имеет свою метку, которая при трансформации одного формата изображения в другой сохраняется в картинке как метаданные. На рисунке 15 представлено преобразованное изображение в формате .png, то есть маска размеченного снимка ОКТ.



Рисунок 15 – Маска размеченного снимка ОКТ в формате .png, где фиолетовым цветом обозначены одиночные друзы, зеленым – сливные друзы

# 3.1 Проведение дополнительной предобработки размеченных изображений

Особенности выбранной для работы архитектуры нейронной сети UNet3+ подразумевают преобразование входных данных до изображения с одинаковым соотношением сторон. В связи с тем, что большинство ОКТ-

снимков в наборе данных OCTDL имеют неравномерное соотношение сторон 800х300 pix, то есть вытянутые горизонтально, было принято решение провести дополнительную предобработку изображений двумя способами: используя функцию cv2.resize и применяя padding.

# 3.1.1 Проведение дополнительной предобработки размеченных изображений с помощью функции cv2.resize

Первый эксперимент по преобразованию соотношения сторон изображений был проведен с использованием встроенной Python-функции cv2.resize с помощью библиотеки OpenCV. Масштабирование картинки в этом случае происходит следующим образом:

1) функция cv2.resize принимает на вход два основных параметра – исходное изображение и желаемый размер выходного изображения, который можно указать через новую ширину и высоту или либо через коэффициент масштабирования;

2) далее функция интерполирует пиксели исходного изображения для создания увеличенного выходного изображения;

3) в конце cv2.resize возвращает увеличенное изображение в виде массива NumPy, который можно дополнительно обработать или отобразить с помощью других функций OpenCV [59].

После применения функции cv2.resize к 160 размеченным изображениям, их разрешение увеличилось до 640х640 ріх. На рисунке 16 представлен результат применения функции cv2.resize к исходному снимку ОКТ.



Рисунок 16 – Исходное ОКТ-изображение (слева), изображение после применения функции сv2.resize (справа)

Из рисунка 16 видно, что целевые объекты, то есть друзы, достаточно сильно искажаются при применении функции cv2.resize к исходному снимку ОКТ, что может повлиять на качество работы модели и итоговые предсказания.

# 3.1.2 Проведение дополнительной предобработки размеченных изображений с помощью Padding

Второй эксперимент был реализован с помощью Padding путем добавления черных пикселей в виде полос снизу и сверху, чтобы изменить размер снимков ОКТ на разрешение 800х800 ріх. В работе был использован Same padding – минимальное заполнение, требуемое в каждом направлении, при котором пиксели добавляется по мере необходимости, чтобы компенсировать перекрытия, когда размер входных данных и размер ядра не идеально совпадают. На рисунке 17 представлен результат применения Padding к исходному снимку ОКТ.



Рисунок 17 – Исходное ОКТ-изображение (слева), изображение после применения Padding (справа)

При сравнении двух экспериментов были получены следующие результаты: дополнительная предобработка размеченных изображений с помощью функции cv2.resize дает следующие метрики качества: Accuracy =  $0,99693 \pm 0,00041$ , Dice Coefficient =  $0,888 \pm 0,00098$ , loss\_function = 0,0106, а предобработка изображений с применением Padding и добавлением черных полос пикселей: Accuracy =  $0,99896 \pm 0,00030$ , Dice Coefficient =  $0,890 \pm 0,00095$ , loss\_function = 0,0054. Из полученных данных следует, что при использовании второго метода обработки данных в работе были получены наилучшие результаты.

#### 3.2 Настройка и обучение нейронной сети UNet3+

Перед обучением нейронной сети UNet3+ была произведена ее настройка и заданы следующие параметры: оптимизатор Adam, в качестве функции потерь была использована Categorical crossentropy (категориальная кросс-энтропия), функция активации ReLu, метриками валидации являются Accuracy и Dice Coefficient, как общий показатель оценки семантической сегментации изображений использован алгоритм IoU, установлено 200 эпох обучения, «batch size» = 16, количество классов n = 2 для 1-го эксперимента и

n = 3 для второго эксперимента, тренировочная и тестовая выборки разделены в процентном соотношении 90:10 соответственно [60].

Алгоритм подготовки данных с последующим обучением нейронной сети UNet3+ представлен на рисунке 18.



Рисунок 18 – Алгоритм обработки данных и обучения нейронной сети

Первым этапом представленного на рисунке 18 алгоритма является разметка набора данных OCTDL с помощью инструмента LabelMe, затем проводится дополнительная предобработка данных с помощью Padding для изменения размера изображений. Далее происходит настройка нейронной сети, выбор оптимальных гиперпараметров, с последующим ее обучением на подготовленном наборе данных. На выходе нейронной сети выдается

предсказанная маска с сегментированными аномалиями по классам заболевания.

Демонстрация работы UNet3+ в данной работе представлена на рисунке 19.



Рисунок 19 – Пример работы UNet3+

Работа модели имеет следующий вид: на вход подаются снимки исходного вида и маска с разметкой заболеваний, на выходе мы получаем предсказанную моделью маску с сегментированными аномалиями по классам: друзы, характеризующие раннюю форму возрастной макулярной дегенерации и сливные друзы, отражающие наличие промежуточной стадии заболевания. Дополнительные примеры работы модели приведены в приложении А.

# 3.3 Решение задачи мультиклассовой сегментации ранней, промежуточной и поздней формы ВМД на снимках ОКТ

В рамках проведения данного исследования была разработана и протестирована мультиклассовой сегментации ранней. модель промежуточной и поздней форм ВМД на основе архитектуры UNet3+ на открытом наборе данных OCTDL. Именно эта модель демонстрирует наилучшие метрики качества работы согласно анализу научной литературы. Для сравнения полученных результатов было обучено две модели: первая из которых предназначена для сегментации и классификации двух классов – одиночные и сливные друзы (160 изображений). Для обучения второй модели, задачей которой является сегментация и классификация одиночных и сливных друз, а также рубцов, была дополнительно проведена разметка и последующая предобработка 76 изображений ОКТ с поздней стадией ВМД. В таблице 4 приведено сравнение результатов качества алгоритмов глубокого обучения.

1 4051	и сравнение результатов ка теетва ал оритмов тлубокого обутения						
N⁰	Статья	Алгоритм	Набор данных	Метрики качества			
1	2	3	4	5			
1	Настоящее исследование (кол-во классов n = 2)	UNet3+	OCTDL	Dice Coefficient = $0,89 \pm 0,00095$ , Accuracy = $0,99896 \pm 0,00030$ , IoU = $0,694 \pm 0,00133$			
2	Настоящее исследование (кол-во классов n = 3)	UNet3+	OCTDL	Dice Coefficient = $0,80 \pm 0,00097$ , Accuracy = $0,99734 \pm 0,00049$ , IoU = $0,670 \pm 0,00137$			

Таблица 4 – Сравнение результатов качества алгоритмов глубокого обучения

1	2	3	4	5
3	Tennakoon, R., и	UNet3+	RETOUCH	Dice Coefficient 0,73
	др. [11]			
4	Kang, S. H., и др.	UNet3+	RETOUCH	Dice Coefficient $= 0,90,$
	[5]			Accuracy $= 0,96$
5	S. Diao, J. Su, C.	UNet3+	UCSD	Dice Coefficient = $0,7751$ ,
	Yang и др. [8]			Accuracy $= 0,9693$

Продолжение таблицы 1

В таблице 4 приведены метрики качества работы моделей нейронной сети, полученные в настоящем исследовании, с доверительными интервалами (ДИ), рассчитанными по формуле (5):

$$\mathcal{A}\mathcal{H} = \bar{x} \pm z \times (s/\sqrt{n}), \tag{5}$$

где  $\bar{x}$  – выборочное среднее,

z – критическое значение для 95 % уровня достоверности составляет 1,96,

s – среднее квадратическое отклонение,

n – размер выборки.

В работах, основанных на наборе данных RETOUCH, Tennakoon, R. и др. получили модель с использованием модифицированного алгоритма U-Net3+ с коэффициентом Dice Coefficient = 0,73 [11]. Капg, S. H. и др., применяя свою модификацию алгоритма U-Net3+ и используя тот же набор данных, получили модель со значением Accuracy = 0,96 и коэффициентом Dice Coefficient = 0,90 [5]. На другом наборе данных UCSD S. Diao и др. были получены следующие результаты: Dice Coefficient = 0,7751, Accuracy = 0,9693 [8].

В рамках проведенного эксперимента модели на архитектуре UNet3+ показали результаты сравнимые с известными [5, 8, 11]. Для двух классов: средняя IoU алгоритма = 69,4 % и это означает, что модель на 70 % правильно сегментирует аномалии, Accuracy = 99,896 % и это означает, что модель с точностью 99 % правильно классифицирует сегментированные аномалии, Loss\_function = 0,0054. Для трех классов полученные значения оказались ниже и составили: IoU = 67,0 %, Accuracy = 99,7344 %, Loss\_function = 0,0144. Более низкие значения метрик можно объяснить тем, что при увеличении количества классов в модели с двух до трех, ее архитектура усложняется, и из-за малого количества обучающих данных, используемых в данной работе, качество модели нейронной сети ухудшается.

Таким образом, используя известную архитектуру UNet3+ для решения задачи мультиклассовой сегментации, на новом открытом наборе данных OCTDL было достигнуто качество работы алгоритма сравнимое с опубликованными ранее научными исследованиями.

#### ЗАКЛЮЧЕНИЕ

Цель, поставленная в начале настоящей работы, достигнута, все задачи решены. В данной работе была разработана и протестирована модель нейронной сети, решающая задачу мультиклассовой сегментации ранней и промежуточной форм возрастной макулярной дегенерации на снимках оптической когерентной томографии глаза. Подводя итоги по проделанной работе, можно сделать следующие выводы:

1) с помощью инструмента LabelMe размечено 236 снимков (160 ранняя и промежуточная форма ВМД, 76 – поздняя форма ВМД) оптической когерентной томографии в формате .json, посредством использования языка Python и библиотеки OpenCV изображения были преобразованы в формат .png с сохранением размеченных классов, таким образом были получены маски. Для работы с нейронной сетью потребовалась дополнительная предобработка, в связи с тем, что на вход модель принимает изображения с одинаковым соотношением сторон. Было протестировано 2 способа изменения размера изображения – это функция resize и заполнение Padding. При использовании первого метода было изменено разрешение изображений 640х640 пикселей и достигнуты следующие значения для модели, решающей задачу сегментации и классификации по двум классам: Accuracy =  $0,99693 \pm 0,00041$ , Dice Coefficient =  $0,888 \pm 0,00098$ , Loss\_function = 0,0106, при использовании второго: Accuracy =  $0,99896 \pm 0,00030$ , Dice Coefficient =  $0,890 \pm 0,00095$ , Loss function = 0,0054, а разрешение составило  $800 \times 800$  пикселей. Таким образом, метод заполнения Padding продемонстрировал себя лучше, т.к. в первом случае целевые объекты, друзы, достаточно сильно искажаются, что снижает качество работы модели;

2) в ходе выполнения исследования было оценено качество работы двух моделей нейронной сети UNet3+ и получены следующие значения:

– для двух классов: Dice Coefficient = 0,89 ± 0,00095 – попиксельно сравнивает соответствие между прогнозируемой сегментацией и

соответствующей ей достоверностью, Accuracy = 0,99896 ± 0,00030 – указывает точность классификации сегментированных аномалий, loss\_function – 0,0054 – показывает на сколько истина отличается от оцениваемого параметра;

-для трех классов: Dice Coefficient = 0,80 ± 0,00097, Accuracy = 0,99734 ± 0,00049, loss\_function - 0,0144;

3) было проведено сравнение полученных результатов с аналогичными работами в области применения глубокого обучения в диагностике возрастной макулярной дегенерации, из которого следует, что в данном исследовании при меньшем количестве обучающих данных достигнуто сравнимое качество работы UNet3+ при решении задачи мультиклассовой сегментации ранней и промежуточной форм ВМД на новом открытом наборе данных ОСТDL.

Исследование, представленное в данной работе, заняло 1 место и было принято к публикации в сборнике всероссийской научной студенческой конференции «ИНТЕР – Информационные технологии и радиоэлектроника». Полученные результаты были апробированы на постерной секции в рамках международной конференции 2024 USBEREIT (2024 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology).

Таким образом, данное исследование позволило решить актуальную на сегодняшний день проблему – автоматизацию диагностики ранней стадии офтальмологических заболеваний на снимках оптической когерентной томографии и повышения точности диагностики данного заболевания с использованием глубокого обучения. Это, в свою очередь, позволяет сократить время, затрачиваемое врачом на интерпретацию снимков, снижает риск субъективной оценки снимков оптической когерентной томографии специалистами, которая может приводить к плохой воспроизводимости и низкому качеству диагностики, а также метод, используемый в настоящей работе, способствует обнаружению заболевания на ранних стадиях. Предложенный способ в дальнейшем открывает возможности применения

модели нейронной сети в системе поддержки принятия решения для средств медицинской диагностики в области офтальмологии за счет увеличения списка размеченных снимков и последующего обучения модели.

#### СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

 Оценка эпидемиологической ситуации и доступности медицинской помощи пациентам с заболеваниями сетчатки глаза в российской федерации / И. В. Ларичева, И. М. Сон, Е. С. Ястребова [и др.] // Менеджер здравоохранения. — 2020. — № 10. — С. 26-36.

 2. Blindness and vision impairment : World Health Organization : офиц. сайт.
 — URL: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visualimpairment (дата обращения: 14.02.2024).

3. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis / W. L. Wong, X. Su, X. Li [et al.] // The Lancet Global Health. — 2014. — V. 2. — No. 2. — P. e106-e116.

4. Topic T.15561 | Part of Topic Cluster TC.33 - Eye; Optical Coherence Tomography; Macular Degeneration : сайт. — URL: https://www.scival.com/trends/summary?uri=Topic/15561 (дата обращения: 10.02.2024).

5. Deep neural networks for the detection and segmentation of the retinal fluid in OCT images / S. H. Kang, H. S. Park, J. Jang, K. Jeon // Proc. MICCAI Retinal OCT Fluid Challenge (RETOUCH). — 2017. — P. 9-14.

6. Deep learning architectures analysis for age-related macular degeneration segmentation on optical coherence tomography scans / K. Alsaih, M.Z. Yusoff, T.B. Tang [et al.] // Computer Methods and Programs in Biomedicine. — 2020. — V. 195. — No. 105566.

7. Multi-scale convolutional neural network for automated AMD classification using retinal OCT images / S. Sotoudeh-Paima, A. Jodeiri, F. Hajizadeh, H. Soltanian-Zadeh // Computers in Biology and Medicine. — 2022. — V. 144. — No. 105368.

8. Classification and segmentation of OCT images for age-related macular degeneration based on dual guidance networks / S. Diao, J. Su, C. Yang, W. Zhu // Biomedical Signal Processing and Control. — 2023. — V. 84. — No. 104810.

9. IA-net: informative attention convolutional neural network for choroidal neovascularization segmentation in OCT images / X. Xi, X. Meng, Z. Qin [et al.] // Biomed Opt Express. — 2020. — V. 11. — No. 11. — P. 6122–6136.

10. Multiclass segmentation as multitask learning for drusen segmentation in retinal optical coherence tomography : сайт. — URL: https://www.researchgate.net/publication/333866227\_Multiclass\_segmentation\_as \_multitask\_learning\_for\_drusen\_segmentation\_in\_retinal\_optical\_coherence\_tomo graphy (дата обращения: 05.03.2024).

11. Retinal fluid segmentation in OCT images using adversarial loss based convolutional neural networks / R. Tennakoon, A. K. Gostar, R. Hoseinnezhad, A. Bab-Hadiashar // 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). — 2018. — P. 1436–1440.

12. Ивахненко О.И. Возрастная макулярная дегенерация и диабетическое поражение глаз. Социально-экономические аспекты заболеваемости / О. И. Ивахненко, В. В. Нероев, О. В. Зайцева // Вестник офтальмологии. — 2021. — Т. 137. — № 1. — С. 123-129.

13. Российская офтальмология онлайн. 1.3. Рецепторная система (сетчатка) : сайт. — URL: https://eyepress.ru/chapter/1-3-retseptornaya-sistema-setchatka- (дата обращения: 07.03.2024).

14. Прогрессивная офтальмология. Сетчатка // сайт. — URL: https://moscow.3z.ru/dictionary/setchatka/ (дата обращения: 07.03.2024).

15. Бикбов М. М. Этиопатогенез неэкссудативной возрастной макулярной дегенерации (обзор литературы) / М. М. Бикбов, Т. А. Халимов // Acta Biomedica Scientifica. — 2022. — Т. 7. — №. 1. — С. 96-105.

16. Risk Factors for the Incidence of Advanced Age-Related Macular Degeneration in the Age-Related Eye Disease Study (AREDS) AREDS Report No.

19 / T. E. Clemons, R. C. Milton, R. Klein [et al.] // Ophthalmology. — 2005. — V.
112. — No. 4. — P. 533-539.

17. Центр сетчатки глаза. Друзы сетчатки : сайт. — URL: https://retinacenter.ru/articles/setchatki-glaza/ (дата обращения: 15.03.2024).

18. Mendeley Data. OCTDL: Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods : сайт. — URL: https://data.mendeley.com/datasets/sncdhf53xc/4 (дата обращения: 15.03.2024).

19. Бровкина А. Ф. Руководство по клинической офтальмологии / А. Ф. Бровкина. —2-е изд., доп. и перераб. —Москва : МИА. — 2024. — С. 336-340.

20. Распространенность офтальмологических заболеваний в популяционной выборке старше 50 лет / И. В. Мунц, А. О. Диреев, О. Г. Гусаревич [и др.] // Вестник офтальмологии. — 2020. — Т. 136. — №. 3. — С. 106-115.

21. Diabetic retinopathy classification for supervised machine learning algorithms / L. F. Nakayama, L. Z. Ribeiro, M. B. Gonçalves [et al.] // International Journal of Retina and Vitreous. — 2022. — V. 8. — No. 1.

22. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales / C. P. Wilkinson, F. L. Ferris, R. E. Klein [et al.] // Ophthalmology. — 2003. — V. 110. — No. 9. — P. 1677-1682.

23. Global Prevalence and Major Risk Factors of Diabetic Retinopathy / J. W.
Y. Yuy, S. L. Rogers, R. Kawasaki [et al.] // Diabetes Care. — 2012. — V. 35. — No. 3. — P. 556–564.

24. Окклюзия ветви центральной вены сетчатки. Патогенез, прогноз зрительных функций, современные способы лечения / С. Н. Тульцева, В. С. Куликов, И. В. Ширяев, Д. А. Яровой // Офтальмологические ведомости. — 2009. — Т. 2. — №. 2, — С. 30-40.

25. Спектральная оптическая когерентная томография: принципы и возможности метода / А. В. Свирин, Ю. И. Кийко, Б. В. Обруч, А. В. Богомолов // РМЖ. Клиническая офтальмология. — 2009. — Т. 10. — №. 2. — С. 50-53.

26. Клиника микрохирургии глаз им. Академика Святослава Федорова. Ангиотомография глаза : сайт. — URL: https://klinikaglaz.ru/services/issledovaniya-i-diagnostika/angiotomografiya-glaza/ (дата обращения: 25.03.2024).

27. Методы машинного обучения в офтальмологии. Обзор литературы /
Д. Д. Гарри, С. В. Саакян, И. П. Хорошилова-Маслова [и др.] //
Офтальмология. — 2020. — Т. 17. — №. 21. — С. 20-31.

28. Development and validation of a computer-aided diagnostic tool to screen for age-related macular degeneration by optical coherence tomography / P. Serrano-Aguilar, R. Abreu, L. Antón-Canalís [et al.] // British Journal of Ophthalmology. — 2011. — V. 96. — No. 4. — P. 503–507.

29. Adaptation of a support vector machine algorithm for segmentation and visualization of retinal structures in volumetric optical coherence tomography data sets / R. J. Zawadzki, A. R. Fuller, D. F. Wiley, B. Hamann // Journal of Biomedical Optics. — 2007. — V. 12. — No. 4. — P. 041206-1–041206-8.

30. Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema / S. J. Chiu, M. J. Allingham, P. S. Mettu [et al.] // Biomed Opt Express. — 2015. — V. 6. — No. 4. — P. 1172-1194.

31. Chapter 9. Convolutional Networks : сайт. — URL: https://www.deeplearningbook.org/contents/convnets.html (дата обращения: 30.03.2024).

32. Сверточные сети : сайт. — URL: https://flectone.ru/svertochniyeseti.html (дата обращения: 30.03.2024).

33. Python School. CNN : сайт. — URL: https://pythonschool.ru/wiki/convolutional-neural-network/ (дата обращения: 03.04.2024).

34. What is the role of 'Flatten' in Keras? : сайт. — URL: https://www.geeksforgeeks.org/what-is-the-role-of-flatten-in-keras./ (дата обращения: 07.04.2024).

35. Основные функции активации в нейронных сетях : сайт. — URL: https://www.yourtodo.ru/posts/osnovnyie-funktsii-aktivatsii-v-nejronnyih-setyah/ (дата обращения: 07.04.2024).

36. Ronneberger O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer. T. Brox // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. — 2015. — V. 9351. — P. 234–241.

37. UNet++: A Nested U-Net Architecture for Medical Image Segmentation /
Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, J. Liang // Deep Learning in Medical
Image Analysis and Multimodal Learning for Clinical Decision Support. — 2018.
— V. 11045. — P. 3-11.

38. UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation : сайт. — URL: https://sci-hub.ru/10.1109/icassp40776.2020.9053405 (дата обращения: 17.04.2024).

39. All the segmentation metrics! : сайт — URL: https://www.kaggle.com/code/yassinealouini/all-the-segmentation-metrics (дата обращения: 17.04.2024).

40. Гришков Д. Ю., Аусилова Н. М. Язык высокого уровня программирования python. Наука и реальность / Д. Ю. Гришков, Н. М. Аусилова // Science & Reality. — 2022. — Т. 9. — №. 1. — С. 114-117.

41. Keras : сайт. — URL: https://blog.skillfactory.ru/glossary/keras/ (дата обращения: 27.04.2024).

42. Что такое NumPy и зачем он нужен : сайт. — URL: https://sky.pro/media/chto-takoe-numpy-i-zachem-on-nuzhen/ (дата обращения: 27.04.2024).

43. OpenCV : сайт. — URL: https://blog.skillfactory.ru/glossary/opencv/ (дата обращения: 27.04.2024).

44. Visualization with Python : сайт. — URL: https://matplotlib.org/ (дата обращения: 29.04.2024).

45. Introduction to Matplotlib : сайт. — URL: https://www.geeksforgeeks.org/python-introduction-matplotlib/ (дата обращения: 29.04.2024).

46. Google Colaboratory : офиц. сайт. — URL: https://colab.google/ (дата обращения: 29.04.2024).

47. Что такое Google Colab и кому он нужен : сайт — URL: https://blog.skillfactory.ru/chto-takoe-google-colaboratory-i-komu-on-nuzhen/ (дата обращения: 29.04.2024).

48. Kaggle : офиц. сайт. — URL: https://www.kaggle.com/ (дата обращения: 30.04.2024).

49. Что такое Kaggle и зачем он дата-сайентисту? : сайт. — URL:https://blog.skillfactory.ru/rabota-s-kaggle-kak-i-s-chego-nachat/обращения: 30.04.2024).

50. Карпова А. Е. Kaggle - платформа для анализа данных / А. Е. Карпова // Вестник магистратуры. — 2018. — Т. 87. — №. 12-4. — С. 48-49.

51. Multivendor Spectral-Domain Optical Coherence Tomography Dataset, Observer Annotation Performance Evaluation, and Standardized Evaluation Framework for Intraretinal Cystoid Fluid Segmentation // J. Wu, A. Philip, D. Podkowinsky [et al.] // Journal of Ophtalmology. — 2016. — V. 2016. — No. 3898750. — P. 1-8.

52. GABNet: global attention block for retinal OCT disease classification / X. Huang, Z. Ai, H. Wang, C. She // Frontiers in neuroscience. — 2023. — V. 17. — No. 1143422. — P. 1-19.

53.OCTID: Optical coherence tomography image database / P. Gholami, P. Roy, M. K. Parthasarathy, V. Lakshminarayanan // Computers & Electrical Engineering. — 2020. — V. 81. — No. 106532. — P. 1-7.

54. OCTDL: Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods / M. Kulyabin., A. Zhdanov, A. Nikiforova [et al.] // Scientific Data. — 2024. — V. 11. — P. 1-10.

55. Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning / T. Schlegl,S.M. Waldstein, H. Bogunovic, F. Endstraßer [et al.] // Ophthalmology. — 2018. — V. 125. — No.4. — P. 549–558.

56. Deep learning approach for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography / F. G. Venhuizen, B. V. Ginneken, B. Liefers [et al.] // Biomedical Optics EXPRESS. — 2018. — V. 9. — No. 4. — P. 15-45.

57. Attention multi-scale network for pigment epithelial detachment segmentation in OCT images / D. Bao, X. Cheng, W. Zhu [et al.] // Proceedings of the SPIE, Medical Imaging 2020: Image Processing. — 2020. — V. 11313. — P. 793-798.

58/ Labelme : сайт. — URL: https://github.com/labelmeai/labelme (дата обращения: 05.05.2024).

59. OpenCV cv2.resize() Function : сайт. — URL: https://www.scaler.com/topics/cv2-resize/ (дата обращения: 05.05.2024).

60. Красильникова Ю.С. Мультиклассовая сегментация друз на снимках оптической когерентной томографии с использованием алгоритмов глубокого обучения / Ю. С. Красильникова, А. Д. Першин // ИНТЕР – Информационные технологии и радиоэлектроника. — 2024. — С. 110-114.

### ПРИЛОЖЕНИЕ А (справочное) Примеры работы модели UNet3+



Рисунок А.1 – Пример работы UNet3+

### Продолжение ПРИЛОЖЕНИЯ А

Маска



### Исходное изображение ОКТ



аномалиями по классам заболевания: друзы (зеленый цвет), сливные друзы (красный цвет)

### Продолжение ПРИЛОЖЕНИЯ А



Исходное изображение ОКТ





Предсказанная маска с сегментированными аномалиями по классам заболевания: рубцы (коричневый цвет)



Рисунок А.3 – Пример работы UNet3+

Маска