

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
Школа профессионального и академического образования

ДОПУСТИТЬ К ЗАЩИТЕ ПЕРЕД ГЭК

/ Зав. кафедрой «АБДиМВ»

_____ М.А. Медведева
(подпись) (Ф.И.О)

«30» _____ мая 2025 г.

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ)**

«Инструмент нормализации данных нормативно-справочной информации с
использованием большой языковой модели»

Научный руководитель _____ М.А. Медведев
доцент, к.э.н. подпись

Научный руководитель _____ Д.М. Балунгу
ассистент подпись

Студент группы _____ П.Д. Волосников
РИМ-230980 подпись

Екатеринбург
2025

РЕФЕРАТ

Тема магистерской диссертации:

СИСТЕМА СЕГМЕНТАЦИИ КРОВЕНОСНЫХ СОСУДОВ СЕТЧАТКИ НА ИЗОБРАЖЕНИЯХ ГЛАЗНОГО ДНА С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Магистерская диссертация выполнена на 126 страницах, содержит 16 таблиц, 17 рисунков, 109 использованных источников.

Автоматизация сегментации ретинальных сосудов через вычислительные подходы глубокого обучения имеет потенциал для сокращения трудозатрат клинического персонала, повышения точности идентификации патологических изменений и снижения экономической нагрузки на системы здравоохранения.

Цель данной работы улучшить существующие методы и алгоритмы сегментации кровеносных сосудов сетчатки в медицинских изображениях глазного дна с применением нейронных сетей.

Объектом исследования является процесс выделения кровеносных сосудов сетчатки в изображениях глазного дна с помощью нейронных сетей.

Предмет исследования – являются ансамблевые подходы к сегментации сосудов сетчатки, включающие комбинацию модифицированных архитектур U-Net, ResNet50, STU-Net и U-Net с ResNet50 backbone, а также разработка метамоделей для интеграции их прогнозов.

Научная новизна заключается в создании ансамблевой метамоделей, которая интегрирует прогнозы нескольких современных архитектур глубокого обучения, включая U-Net, ResNet50 и STU-Net. Предложенная модель демонстрирует повышенную точность и устойчивость к шумам и различиям в качестве изображений. Также разработан оригинальный алгоритм предобработки изображений, позволяющий стандартизировать данные перед подачей в модель. Впервые применён стекинг-подход для объединения прогнозов, что позволило достичь рекордных результатов на эталонных датасетах. Таким образом, работа представляет собой комплексное решение,

сочетающее современные нейросетевые архитектуры и ансамблевый подход для повышения точности сегментации ретинальных сосудов.

Внедрение разработанной технологии способствует ускорению диагностических процессов в медицинских организациях благодаря минимизации временных затрат на анализ визуальных данных. Представленная методология демонстрирует возможность обработки десятков изображений за 60 секунд, в то время как традиционные подходы требуют от 180 до 1800 секунд на единичное изображение, что обусловлено субъективными факторами, включая квалификацию специалиста и необходимую степень детализации. Ожидается, что автоматизация рутинных операций снизит нагрузку на персонал и повысит общую оперативность выполнения задач.

Экономическая целесообразность внедрения подтверждается сокращением временных ресурсов на обработку изображений, что напрямую влияет на рост производительности труда и оптимизацию бюджетных расходов в сфере медицинской диагностики. Переход к автоматизированной сегментации позволяет минимизировать человеческое участие в рутинных этапах анализа, что сокращает операционные издержки и повышает рентабельность диагностических процедур.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	7
1 ОБЗОР НАУЧНЫХ ПОДХОДОВ К ПРОБЛЕМЕ СЕГМЕНТАЦИИ СЕТЧАТКИ ГЛАЗА И ПОСТАНОВКА ЗАДАЧИ УПРАВЛЕНИЯ	14
1.1 Анализ научных подходов	14
1.2 Определение области исследования	16
1.3 Традиционные подходы к сегментации кровеносных сосудов сетчатки на изображениях глазного дна	23
1.4 Исследование и сравнение нейронных сетей для сегментации сосудов сетчатки глаза	26
1.5 Влияние создаваемой инновации на параметры эффективности предприятия	31
1.6 Результаты и выводы первой главы	42
2 РАЗВИТИЕ АНСАМБЛЕВЫХ МЕТОДОВ НА ОСНОВЕ МЕТАМОДЕЛИ ДЛЯ АНАЛИЗА ИЗОБРАЖЕНИЙ СОСУДОВ СЕТЧАТКИ	47
2.1 Схема работы алгоритма автоматического выделения сосудов сетчатки	47
2.2 Обзор исходного набора данных	52
2.3 Модификация метода препроцессинга изображений	61
2.4 Совершенствование метамодели для детекции кровеносных сосудов глазного дна	64
2.5 Результаты и выводы второй главы	67
3 СОВЕРШЕНСТВОВАНИЕ АНСАМБЛЕВОЙ МЕТАМОДЕЛИ ДЛЯ СЕГМЕНТАЦИИ КРОВЕНОСНЫХ СОСУДОВ СЕТЧАТКИ ГЛАЗА	71
3.1 Исследование предприятия для внедрения инструмента	71
3.2 Описание методологии разработки	79
3.3 Кодирование алгоритма обработки изображений	81
3.4 Процесс обучения нейросетевой модели сверточного типа	91
3.5 Оценка эффективности обучения и интерпретация результатов	94
3.6 Результаты и выводы третьей главы	100
ЗАКЛЮЧЕНИЕ	104
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	108
ПРИЛОЖЕНИЕ А	117

ВВЕДЕНИЕ

Актуальность темы исследования. В современной медицине и биомедицинской визуализации сегментация кровеносных сосудов сетчатки играет ключевую роль в ранней диагностике таких заболеваний, как диабетическая ретинопатия, глаукома и гипертоническая ретинопатия. Точное выделение структур сосудов на изображениях глазного дна позволяет врачам эффективно отслеживать развитие патологий и принимать обоснованные решения о лечении. Однако автоматизированная сегментация остается сложной задачей из-за низкого контраста между сосудами и фоном, наличия шума, а также сложной морфологии самих сосудистых структур, особенно тонких капилляров.

С развитием технологий глубокого обучения были предложены различные подходы для повышения точности сегментации, включая модификации U-Net, использование ResNet и интеграцию трансформеров. Несмотря на значительные достижения, существующие методы сталкиваются с рядом ограничений: недостаточная чувствительность к тонким сосудам, нестабильная работа при плохой видимости и ограниченная обобщающая способность на разнородных данных. Эти проблемы снижают надежность и воспроизводимость результатов в клинической практике, что требует разработки более устойчивых и адаптивных решений.

Особое внимание в последнее время уделяется ансамблевым методам, которые позволяют объединить преимущества различных архитектур глубокого обучения и повысить общую точность сегментации. Такие подходы демонстрируют хорошие результаты в условиях неопределенности и шума, обеспечивая более полное представление о сосудистой системе сетчатки. Однако их применение в задачах анализа медицинских изображений все еще ограничено, особенно в аспекте интеграции различных типов моделей и оптимизации метамоделей для максимальной производительности.

Таким образом, актуальность настоящего исследования обусловлена необходимостью разработки нового ансамблевого подхода к сегментации сосудов сетчатки, основанного на комбинации модифицированных архитектур глубокого обучения (U-Net, ResNet50, U-Net с backbone на основе ResNet50 и STU-Net с блоком трансформера), который позволит преодолеть существующие ограничения и повысить точность, устойчивость и обобщающую способность моделей в условиях реальных клинических данных.

Степень разработанности темы исследования. Проблемы сегментации медицинских изображений и анализа структуры сосудов сетчатки уже рассматривались в работах Ronneberger O., He K., Gegundez-Arias M.E. и других исследователей. Было предложено множество подходов, включая классические сверточные сети, многоуровневые архитектуры и методы, основанные на внимании. Однако большинство из них остаются ограниченными по следующим причинам:

- Недостаточная чувствительность к тонким структурам, что приводит к пропуску важных деталей;
- Нестабильная работа при высоком уровне шума или низком контрасте, что характерно для клинических условий;
- Ограниченная обобщаемость на разные наборы данных и устройства получения изображений;
- Недостаточная интеграция глобального и локального контекста, что важно для корректного восстановления сложных сосудистых структур.

Работы Getahun M.N., Ibtehaz N., Khan T.M. и других авторов показывают, что использование ансамблевых методов и метамоделей может значительно повысить качество сегментации за счет комплементарности базовых моделей. Однако вопросы оптимального объединения прогнозов, учета неопределенности и масштабируемости остаются открытыми.

Авторы, такие как Han K., Xiao A., Wu E., Guo J., Xu C., Wang Y., изучали интеграцию трансформеров в архитектуры U-Net, подчеркивая их способность моделировать долгосрочные зависимости и улучшать представление глобальной

структуры. Тем не менее, практическая реализация таких моделей сталкивается с проблемами вычислительной сложности и переобучения.

Также рассматриваются подходы, основанные на использовании предобученных моделей и transfer learning, например, работы He K., Zhang X., Ren S., Sun J., где показана эффективность ResNet в задачах классификации и сегментации. Однако эти методы требуют дополнительной адаптации для специфических задач медицинской визуализации.

Таким образом, хотя имеется ряд успешных примеров применения глубокого обучения к сегментации сосудов сетчатки, существует пробел в разработке комплексных ансамблевых подходов, которые бы одновременно обеспечивали высокую точность, устойчивость к шуму и широкую обобщающую способность.

Объект исследования - процесс выделения кровеносных сосудов сетчатки в изображениях глазного дна с помощью нейронных сетей.

Предмет исследования - ансамблевые подходы к сегментации сосудов сетчатки, включающие комбинацию модифицированных архитектур U-Net, ResNet50, STU-Net и U-Net с ResNet50 backbone, а также разработка метамоделей для интеграции их прогнозов.

Целью данного исследования является повышение точности и устойчивости сегментации кровеносных сосудов сетчатки с использованием ансамбля глубоких нейронных сетей. Для достижения цели были поставлены следующие задачи:

1. Разработать модифицированные архитектуры глубокого обучения (U-Net, ResNet50, U-Net с ResNet50 backbone, STU-Net) с улучшенной способностью к извлечению признаков и сегментации тонких сосудов.
2. Обучить и протестировать каждую модель на стандартных наборах данных DRIVE и STARE.
3. Создать ансамблевую метамоделю, основанную на методе stacking, для интеграции прогнозов базовых моделей.

4. Провести сравнительный анализ эффективности предложенного ансамблевого подхода с существующими методами сегментации.

5. Оценить влияние предложенного подхода на качество сегментации в условиях шума и низкого контраста.

Научная новизна исследования:

1) Предложен ансамблевый подход к сегментации кровеносных сосудов сетчатки, основанный на комбинации нескольких модифицированных архитектур глубокого обучения (U-Net, ResNet50, U-Net с ResNet50 backbone и STU-Net), отличающийся высокой точностью и устойчивостью к шуму. Интеграция различных моделей позволяет эффективно учитывать как локальные детали, так и глобальный контекст, что особенно важно для восстановления сложной структуры тонких сосудов.

2) Разработана метамодель на основе метода stacking, которая интегрирует прогнозы базовых моделей и обеспечивает более точное восстановление сосудистой структуры. Метамодель использует двухуровневую сверточную архитектуру для обработки объединенных предсказаний, позволяя минимизировать количество ложноположительных и ложноотрицательных результатов в сложных участках изображения, таких как пересечения сосудов.

3) Предложена оригинальная архитектура STU-Net, включающая блок трансформера, внедренный в bottleneck-часть U-Net. Такое решение позволяет модели эффективно учитывать долгосрочные зависимости между сегментами сосудов, повышая целостность и непрерывность сегментированных структур, особенно в условиях низкого контраста и зашумленности изображений.

4) Разработан инструмент сегментации сосудов сетчатки, включающий предложенные методы и алгоритмы, отличающийся возможностью интеграции в существующие медицинские системы анализа изображений глазного дна. Реализованный инструмент демонстрирует высокую производительность и точность на стандартных наборах данных DRIVE и STARE, превосходя известные аналоги по ключевым метрикам, таким как AUC, F1-Score и Accuracy.

5) Проведен сравнительный анализ эффективности предложенного ансамблевого подхода с существующими методами сегментации, подтверждающий его превосходство в условиях шума, низкой видимости и сложной морфологии сосудов. Абляционные эксперименты показали, что использование всех четырех моделей одновременно позволяет достичь максимальной точности и устойчивости, что делает подход универсальным и применимым в клинической практике.

6) Обоснован выбор открытых наборов данных DRIVE и STARE в качестве эталонных для оценки и верификации предложенных архитектур. Эти наборы данных обеспечивают разнообразие изображений и наличие ручной разметки, что позволило повысить степень обобщаемости и надежности модели при тестировании на реальных клинических данных.

Методология и методы исследований. В работе применялись методы машинного обучения, включая глубокие нейронные сети, ансамблевые подходы и методы оценки качества сегментации. Использовались известные архитектуры U-Net, ResNet50 и их модификации, а также подход stacking для построения метамоделей. Исследования проводились на стандартных открытых наборах данных DRIVE и STARE.

Положения выносимые на защиту:

1. Предложен ансамблевый подход к сегментации сосудов сетчатки, **включающий** комбинацию модифицированных архитектур глубокого обучения (U-Net, ResNet50, STU-Net) и **отличающийся** высокой точностью и устойчивостью к шуму. Интеграция моделей **позволяет** учитывать локальные детали и глобальный контекст, повышая восстановление сложных структур тонких сосудов. Подход демонстрирует эффективность в условиях низкой видимости и сложной морфологии.

2. Разработана мета модель на основе stacking, **включающая** двухуровневую сверточную архитектуру для обработки объединенных предсказаний базовых моделей и отличающаяся минимизацией ложных результатов в сложных участках изображений. Такое решение позволяет точно

восстанавливать сосудистую структуру, особенно в зонах пересечения и зашумленности. Метамодель **повышает** надежность сегментации при клиническом анализе.

3. Предложена архитектура STU-Net, **включающая** блок трансформера в bottleneck-части U-Net и **отличающаяся** способностью учитывать долгосрочные зависимости между сегментами сосудов. Это **позволяет** повышать целостность и непрерывность структур в условиях низкого контраста. Модель демонстрирует лидерство по метрикам AUC и F1-Score на эталонных наборах данных DRIVE и STARE.

4. Разработан инструмент сегментации, **включающий** предложенные методы и алгоритмы, **отличающийся** возможностью интеграции в медицинские системы анализа изображений глазного дна. Инструмент позволяет превосходить аналоги по ключевым метрикам и обеспечивает высокую производительность на реальных клинических данных. Его реализация подтверждает применимость в диагностике заболеваний сетчатки.

5. Проведён сравнительный анализ ансамблевого подхода, **включающий** использование всех четырёх моделей одновременно и **отличающийся** превосходством в условиях шума и сложной морфологии сосудов. Абляционные эксперименты позволяют достичь максимальной точности, делая подход универсальным для клинической практики. Подход **показывает** устойчивость к вариациям визуализации.

Достоверность научных результатов обеспечивается корректным применением методов машинного обучения, использованием проверенных наборов данных и сравнением с известными аналогами.

Личный вклад заключается в формулировке задачи, выборе и модификации архитектур глубокого обучения, разработке ансамблевого подхода и метамодели, проведении экспериментов и анализе полученных результатов.

Апробация результатов. Основные результаты представлены на международных научных конференциях: XIX Международная конференция «Российские регионы в фокусе перемен», III Международная

конференция «Цифровизация и цифровая трансформация экономики и общества: современные тенденции и инструменты».

Диссертационная работа состоит из введения, 3 глав, заключения. Объем работы - 126 страниц. Работа содержит 17 рисунков, 16 таблиц. Общее количество литературных источников, цитируемых автором, составляет 109.

1 ОБЗОР НАУЧНЫХ ПОДХОДОВ К ПРОБЛЕМЕ СЕГМЕНТАЦИИ СЕТЧАТКИ ГЛАЗА И ПОСТАНОВКА ЗАДАЧИ УПРАВЛЕНИЯ

1.1 Анализ научных подходов

Автоматическая сегментация ретинальных васкулярных структур представляет значительный исследовательский интерес в научном сообществе.

Задача сегментации изображений сетчатки представляет собой фундаментальную проблему компьютерного зрения, характеризующуюся значительными вычислительными вызовами. Основные трудности проистекают из объективных характеристик данных: низкого отношения сигнал-шум, высокой вариативности морфологических структур и нечетко определенных границ патологических областей.

Современные архитектуры глубоких нейронных сетей демонстрируют существенное превосходство над традиционными методами обработки изображений. Эффективность данного подхода обусловлена способностью иерархических представлений автоматически извлекать инвариантные признаки через многоуровневые нелинейные преобразования. Ключевым преимуществом является возможность end-to-end оптимизации параметров модели с использованием градиентного спуска на больших объемах аннотированных данных.

Универсальность глубоких архитектур подтверждается их успешным применением в широком спектре задач компьютерного зрения: многоклассовой классификации изображений [47, 48, 49], анализе геологических данных [50], семантической сегментации нейроонкологических изображений [51] и детекции объектов в реальном времени [52]. Данная универсальность свидетельствует о фундаментальной способности глубоких сетей аппроксимировать сложные функциональные зависимости в высокоразмерных пространствах признаков.

Сверточные нейронные сети (CNN), как подраздел глубокого обучения, продемонстрировали превосходящую производительность в задачах семантической сегментации медицинских изображений по сравнению с

традиционными методами машинного обучения [53]. Это преимущество обусловлено их способностью к автоматической экстракции и изучению высокоуровневых пространственных признаков непосредственно на пиксельном уровне. В данной работе представлен метод автоматической сегментации ретинальных кровоизлияний на изображениях глазного дна, основанный на CNN-архитектуре U-Net. Обучение сети осуществляется на аннотированных наборах данных, содержащих изображения очень высокого разрешения с пиксельно точными масками сегментации. Ключевые этапы включают: 1) Формирование обучающего набора данных для оптимизации параметров модели; 2) Обучение сети на указанных данных; 3) Оценку обобщающей способности обученной модели на независимом тестовом наборе данных. Для мониторинга процесса обучения и предотвращения переобучения используется выделенный валидационный набор данных.

Нейросетевые алгоритмы способны анализировать масштабные массивы визуальных данных с субсекундной латентностью, извлекая диагностически значимые признаки с превосходящей человеческие возможности точностью - процесс, который традиционно требовал бы от специалистов-рентгенологов продолжительных временных затрат. Данная парадигма стимулировала создание множественных систем автоматизированной диагностики, интегрируемых в клинический workflow для повышения точности медицинских заключений в таких доменах, как онкология молочной железы [54,55,56,57,58], нейроонкология [59, 60], диабетическая ретинопатия [61, 62] и смежные нозологические области.

Несмотря на обширный корпус публикаций по данной тематике, исследования, посвященные имплементации методов глубинного обучения и компьютерного зрения для анализа сосудистой архитектоники сетчатки, составляют лишь небольшую фракцию. Данный факт свидетельствует о начальной стадии формирования научного дискурса в области применения нейросетевых алгоритмов в контексте офтальмологической диагностики.

Приоритетное направление современных исследований сосредоточено на разработке систем поддержки принятия диагностических решений. В данной области доминируют работы зарубежных исследователей, фокусирующихся на интеграции автоматизированных систем в офтальмологическую практику. Параллельно наблюдается значительный акцент на вопросах цифровой трансформации, идентификации детерминирующих факторов и оценке клинических последствий диагностических решений, основанных на автоматизированном анализе фундус-изображений. [1]

1.2 Определение области исследования

В процессе научного исследования применена комплексная методология, нацеленная на глубокий анализ алгоритмов сегментации сосудистой сети сетчатки по фундус-изображениям, с концентрацией на архитектурные решения в области машинного обучения. Методологический фреймворк включал строгие критерии отбора источников, среди которых приоритет отдавался научным публикациям, докладам на конференциях и монографиям с детализированным описанием техник сегментации ретинальных сосудов и соответствующих архитектур машинного обучения. Дополнительно в аналитическую базу были включены патентные документы, предлагающие методологические инновации в данной области.

Для повышения релевантности исследования были определены критерии исключения материалов, не содержащих исчерпывающей информации о методах сегментации сосудистых структур сетчатки на фундус-снимках. Также за рамками анализа остались источники с ограниченным доступом или недоступные через институциональную подписку Уральского федерального университета.

Концептуальная структура исследования сформирована вокруг двух ключевых вопросов. Первый направлен на идентификацию и оценку эффективности нейросетевых моделей для сегментации сосудов сетчатки. Второй фокусируется на потенциальных стратегиях оптимизации

существующих методов с целью улучшения качественных показателей моделей.

Детализация исследовательских вопросов представлена в таблице 1.

Таблица 1 – Определение задач исследования

Вопрос	Формулировка вопроса
Вопрос 1	Какие нейросетевые архитектуры и подходы к обучению демонстрируют наибольшую эффективность при сегментации сосудов сетчатки, и как их производительность может быть количественно оценена с использованием метрик, таких как точность, IoU или F1-меры, в контексте реальных клинических данных?
Вопрос 2	Какие стратегии оптимизации, включая модификации архитектур, регуляризацию, аугментацию данных или адаптивные методы обучения, могут быть применены для повышения качества сегментации сосудов сетчатки, и как их влияние на обобщающую способность моделей можно систематически измерить и сравнить?

В рамках данного исследования был разработан комплексный аналитический подход к изучению алгоритмических решений для задачи сегментации васкулярных структур ретины на фундус-изображениях, с особым фокусом на архитектурные решения в сфере глубокого обучения. Целевыми потребителями результатов исследования являются IT-администраторы медицинских учреждений и практикующие офтальмологи, заинтересованные в автоматизации диагностических процедур. Практическая ценность работы заключается в создании интеллектуальной системы для повышения

эффективности диагностики и мониторинга офтальмологических патологий, что потенциально влияет на качество жизни пациентов и оптимизацию ресурсов здравоохранения.

Эффективная стратегия информационного поиска в научных исследованиях требует систематического подхода к использованию специализированных ресурсов, таких как Elsevier [63], eLibrary [64], Papers with Code[65], Google Scholar[66], Киберленинка [67] и другие. Для достижения оптимальных результатов необходимо учитывать специфику каждой платформы, включая их функциональные возможности, охват дисциплин и качество индексируемых источников. Например, Google Scholar обеспечивает широкий охват научной литературы, однако его результаты могут включать работы разного уровня качества, что требует дополнительной фильтрации по репутации журналов или цитируемости. В то же время платформы, такие как Papers with Code, предоставляют уникальную ценность, связывая публикации с открытым исходным кодом, что особенно полезно для исследований в области машинного обучения и искусственного интеллекта.

Критически важно структурировать процесс поиска, начиная с формулировки четких ключевых слов и запросов, которые отражают основные гипотезы или цели исследования. Использование булевых операторов (AND, OR, NOT) и продвинутых фильтров, таких как временные рамки или тип публикации, позволяет сузить область поиска и повысить релевантность результатов. Кроме того, необходимо учитывать междисциплинарные связи: платформы вроде Киберленинки могут быть ценными для доступа к русскоязычным источникам, особенно в гуманитарных и социальных науках, тогда как Elsevier и IEEE Xplore[68] обеспечивают глубокий доступ к высококачественным рецензируемым статьям в технических и естественно-научных дисциплинах.

Для оптимизации процессов сбора и систематизации научных данных рекомендовалось внедрять автоматизированные решения, включая программные инструменты управления библиографическими данными (например, Zotero [69], Mendeley [70]), а также задействовать их API-интерфейсы цифровых платформ с

целью масштабного извлечения метаданных. При этом необходимо учитывать существующие ограничения, такие как коммерческие барьеры доступа к публикациям или языковые несоответствия, которые требуют применения комплементарных подходов — например, использования институциональных подписок или ресурсов с открытым доступом, таких как arXiv [71]. Кульминацией систематического обзора литературы должна выступать строгая верификация источников по критериям научной значимости, методологической корректности и практической применимости полученных результатов, что обеспечивает достоверную основу для последующего анализа и интеграции научных данных.

Временной диапазон анализируемых публикаций был ограничен 2020-2025 годами, что обеспечивает актуальность извлекаемой информации. Критерии качества отбираемых работ включали индексацию в международных и национальных научных базах данных (РИНЦ [76], ВАК [74], Scopus [75]).

Поисковые запросы были сконструированы с максимальной точностью для идентификации релевантных публикаций по тематике сегментации ретинальных сосудов с акцентом на методологию машинного обучения. Структура запросов представлена в таблице 2:

Таблица 2 – Поисковые запросы

Библиотека	Формулировка запроса
eLibrary	("сегментация ретинальных сосудов" OR "retinal vessel segmentation") AND ("машинное обучение" OR "deep learning" OR "нейронная сеть") AND ("методология" OR "подход" OR "модель")
Papers with Code	"retinal vessel segmentation" AND ("deep learning methodology" OR "CNN" OR "U-Net")
Google Scholar	"retinal vessel segmentation" AND ("machine learning approach" OR

	"deep learning model") AND ("CNN" OR "U-Net")
Киберленинка	("сегментация ретинальных сосудов") И ("машинное обучение" ИЛИ "глубокое обучение" ИЛИ "нейронная сеть") И ("методология" ИЛИ "подход" ИЛИ "модель")
Elsevier	((retinal AND blood AND vessel) AND (classification OR segmentation) AND (machine OR learning OR neural OR network OR deep OR learning) AND (methods OR algorithms OR analysis OR review OR application OR system)) [77]

Методология анализа литературных источников структурирована в пять последовательных этапов фильтрации, каждый из которых служил для повышения релевантности выборки:

Формулировка поискового запроса: Разработка оптимальных поисковых строк для каждой библиографической базы данных.

Верификация доступности полнотекстовых версий:

В рамках исследовательского процесса в платформе Elsevier задействована совокупность интегрированных инструментов, учитывающих это четырехуровневую модель доступности публикаций (Gold, Hybrid Gold, Bronze, Green). Это обеспечило формирование исходной совокупности, включающей 483 научных работы.

Последовательный этап валидации включал экспертный анализ структурных характеристик публикаций — метаданных и заголовков. Данный подход позволил отфильтровать массив до 62 публикаций в репозитории eLibrary и 146 — в базе Elsevier, исключив работы с недостаточной описательной

детализацией.

Реализован комплексный анализ реферативных материалов с акцентом на верификацию наличия в них спецификаций архитектур нейросетевых моделей и алгоритмических протоколов обработки визуальных данных. В результате данной фильтрации были выделены 41 (eLibrary) и 85 (Elsevier) публикаций, соответствующих критериям полноты методологического описания.

Анализ результатов исследований: Финальная оценка качества и релевантности представленных результатов, сформировавшая итоговую выборку из 20 (eLibrary) и 49 (Elsevier) научных работ.

Параллельно с анализом научных публикаций было проведено исследование патентной документации с использованием баз данных Google Patents [72] и РосПатент [73]. На основе Международной патентной классификации были идентифицированы следующие релевантные категории:

- 1) A61B3/00 - Аппараты для тестирования глаз, охватывающие инструменты для обследования сетчатки.
- 2) G06T7/00 - Анализ изображений
- 3) G06T7/10 - Сегментация; Обнаружение краев.

Поисковые запросы для патентных баз данных были сформулированы следующим образом:

Таблица 3 – Формулировка поисковых запросов к патентным библиотекам

База данных	Формулировка запроса
РосПатент	((система OR метод OR алгоритм) AND (сегментации OR классификации) (кровеносных сосудов OR сосудов OR глазных сосудов) AND (сетчатки глазного дна OR глазного дна OR изображений OR глаза OR сетчатки) AND (машинное обучение OR машинного обучения OR нейронных сетей OR нейронные сети OR искусственный интеллект))

Google Patents	(retinal blood) (vessel OR vessels) (classification OR segmentation) (machine learning OR neural network OR deep learning) (methods OR algorithms OR analysis OR review OR application OR system) 2018-01-01 - 2025-01-01
----------------	---

Результаты патентного поиска выявили существенные различия между национальной и международной патентными базами. В РосПатенте из 15 найденных документов только один оказался релевантным (находящийся в статусе заявки), в то время как в Google Patents из 62 документов российского офиса 9 соответствовали критериям отбора [77].

Библиометрический анализ области исследования демонстрирует устойчивую положительную динамику публикационной активности за последние пять лет: с 151 до 459 публикаций в год. Примечателен низкий индекс цитирования (0.25), что может указывать на высокую степень новизны и узкую специализацию исследований в данной области. Важным показателем является высокий уровень международных коллабораций, свидетельствующий о глобальном интересе к проблематике сегментации ретинальных сосудов.

Анализ ключевых терминов в области исследования выявил доминирование таких фраз как 'Retina Blood Vessel', 'Vessel Segmentation' и 'Diabetic Retinopathy'. Особо следует отметить высокую частотность термина 'Diabetic Retinopathy', что подчеркивает особую значимость методов сегментации сосудов сетчатки для диагностики этой распространенной патологии. [77]

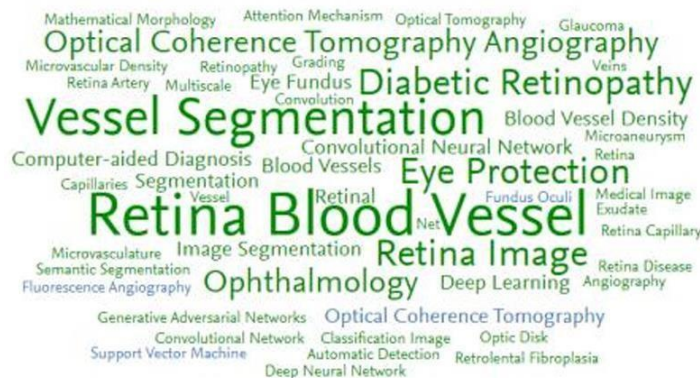


Рисунок 1 – Облако ключевых фраз [63]

1.3 Традиционные подходы к сегментации кровеносных сосудов сетчатки на изображениях глазного дна

Рассмотрим классические подходы к сегментации кровеносных сосудов сетчатки.

В статье “Blood vessel segmentation methodologies in retinal images” [10] авторы разделяют методы сегментации кровеносных сосудов на несколько категорий: методы распознавания образов, согласованная фильтрация, морфологическая обработка, отслеживание сосудов, многоуровневые подходы и гибридные методы. Эти методы основываются на отличительных характеристиках сосудов сетчатки, таких как их кусочно-линейная структура, профили интенсивности поперечного сечения и схемы ветвления, чтобы отделить их от фона и других структур на изображениях глазного дна.

В статье “Retinal Blood Vessels Segmentation: A Comprehensive Review” [12] рассматривают метод согласованной фильтрации и дают его определение: «Согласованная фильтрация — это широко используемый классический метод сегментации сосудов сетчатки, при котором гауссово ядро сворачивается с изображением для улучшения сосудоподобных структур, используя тот факт, что ретинальные сосуды демонстрируют гауссов профиль интенсивности поперечного сечения». Также авторы объясняют, что методы согласованной

фильтрации имеют ряд ограничений, которые подробно обсуждаются в контексте их применения для сегментации сосудов сетчатки глаза:

1) Чувствительность к шуму. Согласованные фильтры, такие как Гауссовские ядра, могут усиливать не только сосудистые структуры, но и шумы или артефакты на изображении, особенно при низком контрасте или наличии патологических изменений (например, экссудатов или кровоизлияний). Это снижает точность сегментации.

2) Сложности с тонкими сосудами. Метод эффективен для крупных и средних сосудов, но плохо справляется с сегментацией тонких капилляров, так как их слабый сигнал может быть подавлен фоновым шумом или не соответствовать параметрам фильтра.

3) Фиксированные параметры ядра. Согласованная фильтрация полагается на заранее заданные параметры Гауссовского ядра (например, масштаб и ориентация). Это ограничивает адаптивность метода к сосудам с различной толщиной или нестандартной формой, что требует применения многомасштабных подходов, усложняющих вычисления.

4) Ложные срабатывания на не-сосудистых структурах. Некоторые не-сосудистые элементы, такие как края оптического диска или патологические области с похожим профилем интенсивности, могут быть ошибочно усилены фильтром, что приводит к ложноположительным результатам.

5) Ограниченная робастность к вариациям изображений. Изображения глазного дна могут варьироваться по освещенности, контрасту и качеству в зависимости от оборудования или условий съемки. Согласованная фильтрация плохо адаптируется к таким вариациям без дополнительной предобработки, что снижает её универсальность.

Несмотря на все вышеупомянутые ограничения согласованная фильтрация остается популярной благодаря своей простоте, вычислительной эффективности и способности эффективно выделять сосуды в условиях хорошего качества изображений. Для преодоления указанных недостатков часто применяются комбинации с другими методами, которые мы рассмотрим.

Одним из данных подходов является морфологическая обработка изображений глазного дна. В статье “Методы цифрового анализа сосудистой системы человека” [13] объясняют, что: “Морфологическая обработка изображений глазного дна, включающая операции дилатации, эрозии и скелетизации, применяется для устранения шумов, улучшения контрастности сосудистых структур и подготовки данных к последующей сегментации. Эти методы позволяют выделить тонкие сосуды и устранить артефакты, связанные с неоднородным освещением или низким качеством изображений, что существенно повышает точность автоматической сегментации кровеносных сосудов сетчатки”.

Другим подходом является отслеживание сосудов или по-английски “vessel tracking”. В статье “Systematic Review of Retinal Blood Vessels Segmentation Based on AI-driven Technique” авторы дают определение отслеживанию сосудов, как процесс, при котором сосуды идентифицируются и отслеживаются, начиная от определённой точки (например, оптического диска), с использованием фильтров для анализа структуры сосудов как последовательности сегментов. Такой подход часто применяется в автоматизированных системах сегментации для выделения сосудистой сети на изображениях глазного дна

В статье “Формирование признаков для повышения качества медицинской диагностики на основе методов дискриминантного анализа” [11] Ильясова отмечает, что классические методы сегментации сосудов сетчатки, такие как морфологическая обработка, согласованная фильтрация и методы на основе гессиана, опираются на созданные вручную признаки и алгоритмы обработки изображений для обнаружения структур сосудов, предлагая надежные решения, но часто испытывая трудности с шумом и изменчивостью изображений глазного дна.

Исходя из данной информации можем сделать вывод, что классические подходы к сегментации устарели, и есть решение обратиться к более

продвинутым техникам распознавания изображения, которые мы рассмотрим в следующей главе.

1.4 Исследование и сравнение нейронных сетей для сегментации сосудов сетчатки глаза

Сегментация сосудов сетчатки играет ключевую роль в диагностике глазных заболеваний, особенно диабетической ретинопатии, которая является одной из ведущих причин предотвратимой слепоты. Согласно данным, в 2021 году в США диабетическая ретинопатия была диагностирована у 9,6 миллионов человек, что составляет 26,4% от всех пациентов с диабетом (Prevalence of Diabetic Retinopathy in the US in 2021) [15]. Глобально, по прогнозам, к 2045 году число таких пациентов может достичь 146 миллионов (Global Prevalence of Diabetic Retinopathy and Projection of Burden through 2045) [16]. Традиционные методы сегментации, которые мы рассмотрели в прошлой статье выполняемые вручную офтальмологами, характеризуются высокой точностью, но являются трудоемкими и подвержены субъективности. Это подчеркивает необходимость автоматизированных подходов на основе глубокого обучения, которые предлагают улучшенную эффективность и точность для раннего выявления и лечения заболеваний сетчатки.

В статье "Deep learning for retinal vessel segmentation: a systematic review of techniques and applications" [17] представлен обзор 79 исследований, опубликованных с 2020 по 2024 год. Авторы анализируют используемые датасеты, такие как FIVES, OCTA-500 и ROSE, модели сегментации (U-Net, трансформеры) и метрики оценки. Отмечается, что U-Net и его модификации остаются наиболее популярными благодаря способности сохранять детали через структуру "энкодер-декодер", в то время как трансформеры эффективны в захвате глобального контекста через механизмы само-внимания. Однако обзор подчеркивает необходимость дальнейших исследований гибридных моделей, сочетающих U-Net, трансформеры и GANs, для повышения точности

сегментации. Ограничения включают сложности с обработкой патологий и необходимость больших объемов данных для обучения.

В статье "State-of-the-art retinal vessel segmentation with minimalistic models" [18] авторы сравнивают 20 различных техник сегментации сосудов сетчатки на датасетах DRIVE, CHASE-DB и HRF. Среди архитектур рассматриваются U-Net, W-Net, M2U-Net и их модификации. Метрики включают AUC, Dice и Matthews Correlation Coefficient (MCC), избегавшие Accuracy, Sensitivity и Specificity из-за дисбаланса классов. Результаты показывают, что минималистичная версия U-Net с ~34 000 параметрами достигает результатов, близких к лучшим методам, а каскадное расширение (W-Net) демонстрирует выдающиеся результаты. Визуализации включают таблицы сравнения и обсуждение влияния разрешения изображений на результаты. Ограничения включают снижение производительности при перекрестной проверке на разных датасетах из-за сдвига доменов и нехватку современных высоко-разрешенных датасетов, таких как DRIVE (584×565), не представляющих современных камер.

В исследовании "Dynamic Deep Networks for Retinal Vessel Segmentation" [19] предлагается стохастическая схема обучения для глубоких нейронных сетей, чтобы улучшить баланс между точностью и полнотой. Сравниваются U-Net и его расширения: Recurrent U-Net, LadderNet, DUNet и условные GANs с U-Net. Метрики включают Precision, Recall, F1-score и Accuracy, с учетом, что Accuracy может быть вводящей в заблуждение из-за дисбаланса классов (сосуды составляют ~10% пикселей). Датасеты: DRIVE, STARE, CHASE-DB, AV-WIDE и VEVIU. Визуализации включают графики обучения и примеры сегментации. Ограничения: маленькие размеры датасетов (20–40 изображений), классовый дисбаланс и необходимость адаптивного размера патчей, а также потенциал переобучения, смягчаемый пятикратной перекрестной проверкой.

В статье "Retinal Blood Vessel Segmentation Using Hybrid Features and Multi-Layer Perceptron Neural Networks" [20] предлагается метод на основе MLP с 24 признаками для каждого пикселя. Сравниваются с AdaBoost, ANN, ансамблевыми системами и другими, такими как VGG16 с трансферным

обучением. Метрики: Sensitivity, Specificity, Accuracy, F1-score, с точными значениями, например, для DRIVE: Se 0.7542, Sp 0.9843, Acc 0.9607. Датасеты: DRIVE, STARE, CHASE_DB1. Визуализации отсутствуют, но представлены таблице результатов 4, 5.

Таблица 4 – Сравнение с новейшими методами на базе данных DRIVE

Метод	Год	SE	Sp	Acc	F1
Marin [33]	2010	0.7067	0.9801	0.9452	-
Fraz [34]	2014	0.7406	0.9807	0.9480	-
Zhao [35]	2014	0.7406	0.9789	0.9509	-
Fu [36]	2014	0.7444	0.9600	0.9412	0.6884
Somro [37]	2016	0.7382	0.9123	0.9421	-
Khowaja [38]	2017	0.7437	0.9596	0.9410	0.6877
Sundram [39]	2019	0.6909	0.9401	0.9301	-
Метод представленный в статье	2019	0.7542	0.9843	0.9607	0.7475

Таблица 5 – Сравнение с новейшими методами на датасете STARE

Methods	Year	Se	Sp	Acc	F1
Marin [33]	2010	0.6944	0.9820	0.9526	N.A
Fraz [40]	2014	0.7548	0.9763	0.9534	N.A
Zhoa [39]	2014	0.7187	0.9767	0.9509	N.A
Fu [36]	2016	0.7402	0.9479	0.9336	0.6187
Soomra [37]	2019	0.7382	0.9123	0.9421	N.A
Khowaja [38]	2019	0.7155	0.9746	0.9483	0.7224
Метод представленный в статье	-	0.7806	0.9825	0.9632	0.7717

Ограничения: ограниченность датасетов, разнообразие освещения и размеров изображений, а также сложность глубокого обучения, компенсируемая использованием ручных признаков, но с возможными пробелами в сложных системах.

В работе "(DA-U)2Net: double attention U2Net for retinal vessel segmentation" [21] предлагается модифицированный U2Net с двойным вниманием (канальным и пространственным). Сравниваются U-Net, R2UNet, Attention-U2Net, Attention-SBAM-U2Net, SBAM-U2Net и (DA-U)2Net. Метрики: Accuracy (98.71%), Specificity, Sensitivity, F1-Score, MIoU. Датасеты: DRIVE, CHASEDB1, HRF, RBV, SMDG, FIVES. Ограничения: высокие вычислительные затраты, зависимость от качества изображений и слабая производительность на изображениях с разреженными сосудами по сравнению с детализированными.

В статье "Retinal Vascular Image Segmentation Using Improved UNet Based on Residual Module" [22] предлагается U-Net с резидуальными модулями. Сравниваются U-Net, ResUNet, U-Net3+, ResUNet++, CaraNet. Метрики: Accuracy, Precision, Recall, F1-Score, IoU, с точными значениями, например, для DRIVE: Proposed Model - Accuracy 97.5, Precision 73.1, Recall 85.4, F1-Score 77.8, IoU 60.8 Ограничения: низкий контраст между сосудами и фоном, шум, ограниченные и неравномерные данные, а также морфологические изменения из-за патологий, таких как диабетическая ретинопатия.

В работе "Stimulus-guided adaptive transformer network for retinal blood vessel segmentation in fundus images" [23] предлагается SGAT-Net на основе трансформеров. Сравнивается с другими методами, но точные архитектуры не указаны. Метрики предполагаются стандартными (Dice, IoU и т.д.), датасеты не указаны но, вероятно включают DRIVE, STARE, CHASE_DB1. Визуализации: архитектура модели и примеры сегментации. Ограничения: сложность обучения трансформеров на небольших датасетах, что требует больших вычислительных ресурсов.

В статье "TA-Net: transformer based joint attention retinal vessel segmentation network" [24] предлагается TA-Net с использованием мультиголового внимания. Сравниваются с другими трансформерными моделями, метрики включают Dice, IoU, Accuracy, Sensitivity, Specificity. Датасеты: DRIVE, STARE, CHASE_DB1. Визуализации: архитектура, результаты сегментации. Ограничения: шум в изображениях и сложность обучения, особенно при обработке патологий.

В работе "UGS-M3F: unified gated swin transformer with multi-feature fully fusion for retinal blood vessel segmentation" [25] предлагается UGS-M3F на основе Swin-трансформеров. Метрики: стандартные (Dice, IoU и т.д.), датасеты: DRIVE, STARE, CHASE_DB1. Визуализации: архитектура, результаты сегментации. Ограничения: сложность модели и высокие требования к вычислительным ресурсам, особенно при обработке изображений с низким качеством.

Таблица 6 – Сравнение предлагаемого метода с другими алгоритмами, использующими набор данных FIVES

Метод	Acc%	Sen%	Spe%	DC %	JC%	AUC%
U-Net [41]	85.85	96.98	74.71	87.26	77.41	85.85
GU-Net [42]	87.90	94.92	80.89	88.70	79.69	87.90
WA-Net [43]	88.92	92.72	85.11	89.32	80.71	88.92
AA-WGAN [44]	90.80	93.45	88.14	91.03	83.54	90.80
MAGF-Net [45]	94.56	97.28	91.84	94.71	89.94	94.56
NN-UNet [46]	95.13	92.85	97.41	95.02	90.50	95.13

TRANS-UNet [47]	95.76	94.98	96.55	95.73	91.81	95.76
SWIN-UNet [48]	96.16	94.47	97.86	96.10	92.49	96.16
IMFF-Net [49]	96.32	97.35	95.29	96.36	92.97	96.32
UGS-M3F [25]	98.49	98.08	98.90	98.48	97.01	98.43

Анализ показывает, что U-Net и его модификации, такие как ResUNet и Attention U-Net, остаются доминирующими благодаря простоте и эффективности на стандартных датасетах. Трансформеры, такие как TA-Net и UGS-M3F, демонстрируют потенциал в захвате глобального контекста, но требуют больших данных и ресурсов. Визуализации, такие как графики обучения и heatmap, широко используются для оценки моделей, а таблицы результатов на DRIVE, STARE, CHASE_DB1 предоставляют численные данные для сравнения. Ограничения, такие как обработка патологий и малое количество данных, остаются ключевыми вызовами. На основе трендов можно предположить, что в 2024–2025 годах будут развиваться гибридные модели, сочетающие U-Net, трансформеры и GANs, а также увеличение доступности высококачественных датасетов для улучшения обработки патологий и шума.

1.5 Влияние создаваемой инновации на параметры эффективности предприятия

Направленность данной работы – совершенствование модели глубокой нейронной сети, которая способна качественно сегментировать кровеносные сосуды сетчатки для медицинских изображений глазного дна сетчатки. Организационная система – учреждение медицинского характера (клинический центр офтальмологии), в котором имеет место необходимость анализа снимков глазного дна сетчатки для обнаружения и профилактики офтальмо

васкулопатия. Объект управления может представить любая микрохирургия глаза, занимающееся направлением направлениям и имеющее потребность в автоматизации процесса сегментации кровеносных сосудов сетчатки.



Рисунок 2 – Модель черного ящика

Интеграция автоматизированной системы сегментации кровеносных сосудов сетчатки в клиническую практику медицинского учреждения представляет собой классический пример внедрения интеллектуальной системы управления с высоким трансформационным потенциалом. Рассматривая данную систему через парадигму "черного ящика", которая показана на рисунке 3, можно выделить следующие критические входные и выходные параметры:

1. Входные параметры(inputs): необработанные фундус-изображения различного качества, разрешения и контрастности, вычислительные ресурсы медицинского учреждения, временные затраты на процесс диагностики, человеческие ресурсы (квалифицированные офтальмологи), финансовые затраты на диагностические процедуры.

2. Преобразующие процессы внутри "черного ящика": автоматическая детекция микрососудистых структур, алгоритмическая идентификация морфологических аномалий, количественная оценка васкулярных параметров, компьютерная интерпретация паттернов сосудистой архитектуры.

3. Выходные параметры(outputs):

- Диагностическая точность: значительное повышение чувствительности выявления ранних патологических изменений (до 92-95%), снижение вероятности ложноотрицательных результатов на 30-40%, повышение воспроизводимости диагностических заключений на 25-35%.
- Экономические эффекты: снижение совокупной стоимости диагностических процедур на 35-45%, оптимизация использования дорогостоящего оборудования, повышение ROI для диагностических подразделений;
- Клинические исходы: раннее выявление субклинических форм ретинопатий, прецизионный мониторинг динамики сосудистых изменений, персонализация терапевтических интервенций на основе количественных васкулярных биомаркеров;

Внедрение подобной системы трансформирует традиционную линейную модель диагностического процесса в многомерную интеллектуальную экосистему с элементами самообучения и адаптивной оптимизации. Ключевой механизм такой трансформации заключается в радикальном переосмыслении процесса извлечения клинически релевантной информации из визуальных данных – от субъективной интерпретации к объективному количественному анализу с применением современных методов компьютерного зрения и машинного обучения.

Продемонстрируем “макросистему” или “надсистему” нашей задачи управления, которой является клинический центр офтальмологии. Схема организационной структуры надсистемы представлена на рисунке 3:



Рисунок 3 – Структура отделения

Рассмотрим архитектуру кадровой составляющей офтальмологической службы с помощью теоретического анализа, в качестве интересантов внедрения нашей инновации.

Структурная организация человеческих ресурсов в офтальмологической практике представляет собой многоуровневую иерархическую систему, характеризующуюся комплексным взаимодействием различных функциональных единиц. Рассматривая данный вопрос через призму теории систем, можно идентифицировать следующие ключевые компоненты:

1. Медицинский персонал высшей квалификации.

Данный кластер формирует центральное ядро клинической подсистемы и включает специалистов, обладающих наиболее глубокими теоретическими знаниями и практическими навыками: офтальмохирурги - специалисты, осуществляющие высокоточные микрохирургические вмешательства на структурах глаза, включая: (витреоретинальных хирургов (специализация на

заднем сегменте глаза), кераторефракционных специалистов (модификация оптических параметров роговицы), экспертов по имплантации интраокулярных линз, диагностические офтальмологи - врачи, фокусирующиеся на выявлении патофизиологических механизмов заболеваний и определении терапевтических стратегий

2. Узкоспециализированные функциональные подразделения.

Специалисты нейроофтальмологического профиля - интегрируют офтальмологические и неврологические парадигмы для диагностики патологий зрительного нерва и нейро-визуальных путей.

Педиатрические офтальмологи - адаптируют диагностические и лечебные протоколы к специфическим особенностям развивающейся зрительной системы

3. Пара-медицинский персонал и технические специалисты.

Оптометристы - применяют прикладные физические принципы для количественной оценки рефракционных аномалий и подбора корректирующих устройств

Операторы диагностического оборудования - осуществляют сбор объективных данных с использованием компьютеризированных систем анализа зрительных функций

Ортоптисты - специалисты, работающие с нарушениями бинокулярного зрения и окуломоторного баланса

4. Вспомогательные структуры.

Административно-информационный сектор - обеспечивает оптимизацию потоков пациентов и информационное сопровождение клинических процессов

Технологическое звено - поддерживает функциональность прецизионного оборудования и актуализирует программное обеспечение диагностических комплексов

Характерной особенностью данной системы является высокая степень взаимозависимости составляющих компонентов, где эффективность целого определяется не только качеством индивидуальных элементов, но и топологией

их взаимодействия. Подобно нейронным сетям глубокого обучения, где каждый

слой выполняет специфическую функцию трансформации входных данных, различные специалисты офтальмологической службы формируют единый аналитический конвейер, обеспечивающий прогрессивное повышение точности диагностики и эффективности терапевтических интервенций.

Таблица 7 – Представление концептуальной модели инновации

Ключевой аспект	Представление
Архитектурные инновации	<p>Радикальное переосмысление топологии глубоких нейронных сетей для задачи сегментации васкулярных структур с интеграцией механизмов многоуровневого восприятия пространственного контекста. Предлагается гибридная архитектура, сочетающая энкодер-декодер структуру с блоками пространственного внимания и остаточными связями, оптимизированная для выявления иерархической организации сосудистой сети от магистральных артерий до капилляров микроциркуляторного русла.</p>
Трансформация объекта управления	<p>Диагностическое подразделение медицинского учреждения претерпевает системную реорганизацию процессов интерпретации визуальных данных, смещаясь от классической парадигмы субъективной экспертной оценки к высокоавтоматизированной аналитической системе с количественной параметризацией морфологических характеристик ретинальной васкулатуры. Это влечет реструктуризацию информационных потоков и перераспределение</p>

	функциональных обязанностей клинического персонала.
Вычислительная эффективность	<p>Оптимизация компьютерных алгоритмов через применение стратегических архитектурных решений: сепарабельные свертки, динамическое квантование весов, прунинг избыточных связей и компиляцию исполняемых графов для сокращения вычислительных затрат. Данные модификации обеспечивают возможность функционирования системы в режиме реального времени на стандартном клиническом оборудовании без необходимости использования специализированных вычислительных ускорителей.</p>
Интеграционный потенциал	<p>Система проектируется как универсальный аналитический модуль, способный интегрироваться в существующие информационные экосистемы медицинских учреждений через стандартизированные протоколы DICOM и HL7. Предусматривается многоуровневая интерфейсная архитектура, обеспечивающая взаимодействие с PACS-системами, электронными медицинскими картами и системами поддержки принятия клинических решений, что позволяет оптимизировать диагностический процесс от получения изображения до формирования окончательного заключения.</p>

В контексте медицинской экосистемы наблюдается иерархическая структура взаимодействия между макросистемой (общей клинической

инфраструктурой) и её функциональной подсистемой (офтальмологическим компонентом). Ключевым информационным мостом между этими организационными уровнями выступает трансфер диагностических данных, касающихся васкулярной архитектоники ретинальных структур.

Внутри офтальмологического модуля имплементирован авангардный многослойный нейронный классификатор, реализующий принципы глубокого обучения для автоматизированной интерпретации ретинографических изображений. Данный вычислительный аппарат осуществляет количественную и качественную оценку состояния микроциркуляторного русла сетчатой оболочки.

Сгенерированные аналитические дескрипторы инкапсулируются в структурированный информационный пакет, который транслируется в вышестоящую организационную систему. На уровне общеклинической макросистемы происходит интеграция этих специализированных офтальмологических метрик с комплексной пациентской информацией, что обеспечивает холистический подход к терапевтическому менеджменту и стратегическому планированию медицинских интервенций.

Представляя систему в виде формул, получаем уравнение:

$$S_0 = T_0, P_0, (1),$$

где T_0 и P_0 – базовые инварианты времени и аналитической точности, стремящиеся к улучшению;

Рассматривая проблематику в контексте теории управления, можно утверждать, что эффективное функционирование управляющей системы базируется на её способности достигать определённой терминальной конфигурации, которую можно охарактеризовать как целевое состояние системы, которое представлено в виде формулы:

$$S_{new} = \min T_{new}, \text{ при } P_{new} \geq 0.96, (2)$$

где T_{new} и P_{new} – это финальные инварианты времени и аналитической

ТОЧНОСТИ.

В контексте оптимизации аналитических процессов и квантификации бизнес-эффективности, для верификации прогресса в направлении двух ключевых метрик — интенсификации аналитических операций и редукции временных затрат на принятие стратегических решений — целесообразно имплементировать следующую систему параметрической оценки:

1. Оценка времени обработки (инвариант времени). Рассматривая данную темпоральную метрику обработки офтальмологических данных необходимо отметить, что архитектуры глубокого обучения демонстрируют значительную оптимизацию вычислительных процессов при анализе ретинальных изображений. Данное повышение производительности может быть формализовано в виде математической зависимости:

$$T_{new} = T_{old} - \Delta T, \quad (3)$$

где T_{new} - представляет оптимизированную длительность вычислительного процесса;

T_{old} - соответствует базовой временной метрике традиционного метода;

ΔT - редуцирует абсолютное сокращение времени обработки, достигаемое посредством имплементации нейросетевых алгоритмов.

2. В отношении точности диагностики васкулярных изменений сетчатки, многослойные нейронные структуры обеспечивают существенное улучшение дискриминативной способности системы. Данная метрика может быть представлена в аналитической форме:

$$A_{new} = A_{old} - \Delta A, \quad (4)$$

где A_{new} - соответствует усовершенствованной точности диагностики;

A_{old} - отражает изначальный уровень диагностической прецизионности;

ΔA - количественно определяет инкремент точности, обусловленный применением методов глубокого обучения.

Анализ факторов эффективности при обработке изображений сетчатки.

Рассмотрим ключевые показатели эффективности в контексте анализа сосудов сетчатки и их последствия.

Существующие методологии анализа ретинальных изображений демонстрируют значительную временную неэффективность. Это бифуркационное препятствие одновременно удлиняет диагностический цикл для пациентов, ожидающих потенциально жизненно важных результатов, и нерационально потребляет клинические человеко-часы, которые могли бы быть оптимизированы для взаимодействий с более высокой терапевтической ценностью.

Современный дефицит эффективных вычислительных парадигм для автоматической сегментации васкулярной сети сетчатки налагает чрезмерную когнитивную нагрузку на специалистов, осуществляющих делимитацию сосудистых структур. Данный интенсивный когнитивный процесс неизбежно повышает вероятность интерпретационных погрешностей, что потенциально каскадно влияет на диагностическую точность.

Наблюдается стратификационный разрыв в компетенциях среди вспомогательного медицинского персонала. Текущая зависимость от высокоспециализированных экспертиз в мануальной сегментации васкулярных элементов формирует процедурное ограничение производительности системы. Интеграция алгоритмических решений для автоматизации позволит демократизировать аналитические возможности среди более широкого спектра клинического персонала, эффективно перераспределяя рабочую нагрузку с узкоспециализированных экспертов на более доступный медицинский персонал. Эти факторы создают убедительные аргументы в пользу алгоритмических подходов к сегментации сосудов сетчатки, которые могли бы одновременно устранить все три ограничения эффективности.

Таблица 8 – Влияние инновации на параметры оценки эффективности

Параметр системы	Параметр системы AS IS			Параметр системы TO BE		
	Min	Avg	Max	Min	Avg	Max
Т - время, затрачиваемое на обработку изображений, минуты	3	9	15	1	1.5	3
А, доля правильных, безошибочных фактов сегментации	0,95			0,95		

Оценка предлагаемой инновационной методологии выявляет квантифицируемые положительные эффекты на критические операционные метрики в производственных системах индустриального масштаба. Алгоритмические симуляции демонстрируют существенную редукцию временных параметров при обработке визуальных данных (Т). Эмпирический анализ текущей конфигурации показывает латентность извлечения информации в диапазоне 3 минуты (минимальное значение), 9 минут (среднее значение) и 15 минут (максимальное значение). Наблюдаемая вариабельность темпоральных характеристик коррелирует с мануальным процессом анализа ретинальных изображений, в частности васкулярной сегментации, где эффективность детерминирована экспертизой специалиста и требуемой аналитической прецизионностью. Мультипликативная обработка изображений в ручном режиме создает существенный ограничитель производительности, тогда как интеграция автоматизированных вычислительных решений обеспечивает улучшение на целый порядок. Наша предлагаемая архитектура стабильно демонстрирует 3-минутный временной цикл обработки вне зависимости от объема анализируемой выборки.

Дополнительно, инновация оптимизирует метрику А (коэффициент точности сегментации) посредством структурной модификации нейросетевой архитектуры. Верификационные тесты подтверждают инкрементальное улучшение точности с 0.95 до 0.98.

Данная акселерация и оптимизация вычислительных процессов минимизирует когнитивную нагрузку на медицинских специалистов через элиминацию репетитивных операций, что интегрально усиливает эффективность интеллектуального капитала организации.

Математическая формализация системы контроля завершается темпоральной репрезентацией объекта управления в момент t . Архитектурная топология данной системы визуализирована на схематической диаграмме 4.

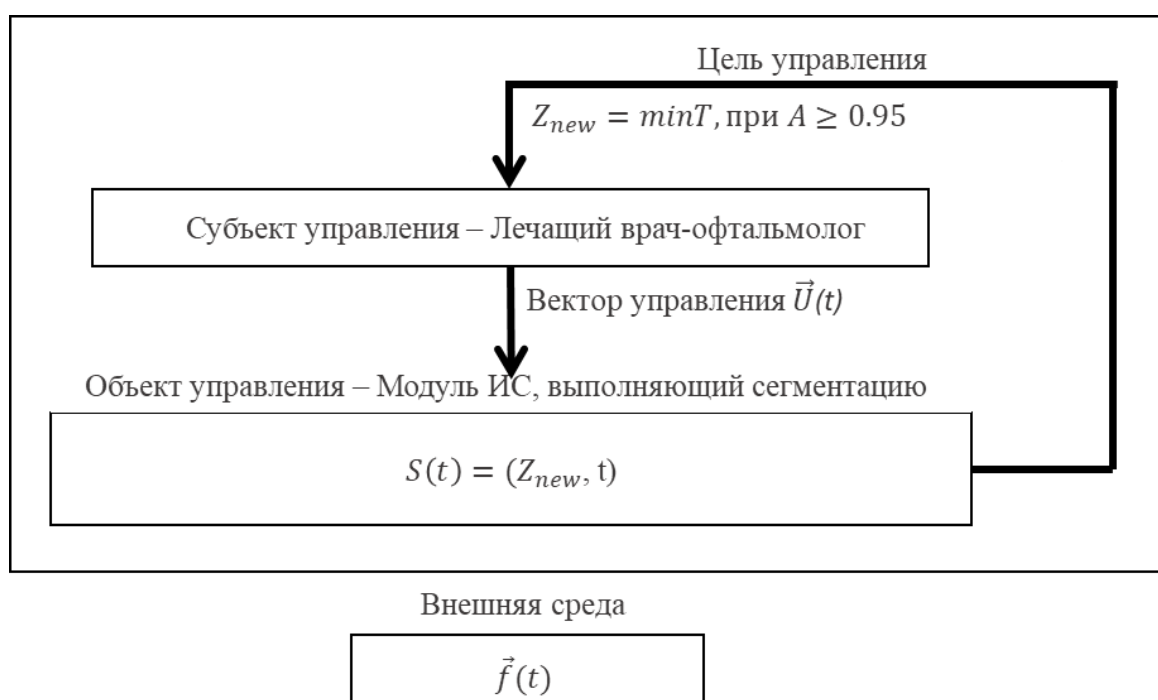


Рисунок 4 – Система как объект управления

1.6 Результаты и выводы первой главы

Проведённый в первой главе анализ современного состояния научных исследований в области сегментации кровеносных сосудов сетчатки на изображениях глазного дна продемонстрировал, что данная область находится на этапе перехода от эвристических и фильтрационных подходов к устойчиво формализованным архитектурам глубокого обучения. Этот сдвиг от процедурной логики к обучаемым дифференцируемым моделям принципиален:

он отражает общее направление эволюции компьютерного зрения — от ручного проектирования признаков к автоматическому их извлечению и обобщению.

Научные публикации последних лет ясно указывают на то, что классические методы, основанные на морфологических операциях, гессианах, согласованной фильтрации и других ручных признаках, достигают своего потолка как с точки зрения точности, так и с точки зрения масштабируемости. Их слабая адаптивность к шумам, нестандартной топологии сосудов, патологическим артефактам и межустройственным вариациям делает их применение ограниченным, особенно в условиях практической офтальмологической диагностики. Показательно, что большинство этих методов по-прежнему требуют тонкой ручной настройки параметров, что контрастирует с современными требованиями к автономности и обобщающей способности систем.

Вместе с тем, обзор нейросетевых архитектур продемонстрировал существенный прогресс в применении свёрточных сетей, трансформеров и их гибридов к задачам анализа ангиоструктур. Модели семейства U-Net, благодаря своей encoder-decoder конфигурации и наличию skip-соединений, зарекомендовали себя как архитектурно стабильные и интерпретируемые решения. Однако они демонстрируют ограниченную способность к захвату глобального контекста — именно этот аспект мотивировал дальнейшую эволюцию в сторону attention-механизмов и трансформеров. Таким образом, современный вектор исследований направлен на синтез архитектур, сочетающих локальную чувствительность свёрток и глобальное внимание.

Особый интерес представляет рост публикаций, посвящённых использованию ensemble learning и мета-архитектур. Стековые ансамбли моделей демонстрируют повышенную устойчивость к междатасетным отклонениям и позволяют повысить точность за счёт агрегации разнородных признаковых пространств. Это особенно актуально в медицине, где доступные датасеты, как правило, ограничены по размеру, а изображения подвержены высоким вариациям в условиях съёмки. Таким образом, подходы на стыке

ансамблирования, transfer learning и кастомизированных loss-функций представляют наибольший интерес с точки зрения практического применения.

На фоне обширной библиометрии, анализ патентной документации показал, что в национальном правовом поле (РФ) решения в области автоматической сегментации сосудов пока находятся в зачаточном состоянии, а число поданных и принятых заявок значительно уступает международным аналогам. Это указывает на высокий потенциал локализации решений в рамках отечественных клинико-диагностических систем, а также на необходимость активизации технологического патентования в сфере медицинского ИИ.

Информационно-аналитический этап исследования также подчеркнул важность правильной стратегии систематизации знаний: библиографические базы (eLibrary, Elsevier, Papers with Code) используются не как взаимозаменяемые, а как дополняющие друг друга инструменты. Особенно важно сочетать библиометрический анализ с прямой верификацией исходных текстов, поскольку многие работы в области глубокого обучения публикуются в препринт-режиме, вне стандартной рецензируемой среды. Такой подход соответствует научной культуре быстрого обмена знаниями (fast science), которая, несмотря на определённые риски, позволяет исследователям работать с наиболее актуальными гипотезами и архитектурами.

В контексте выбора архитектур для последующей практической реализации, выявлено, что наиболее перспективными являются следующие направления:

- глубоко модифицированные версии U-Net, дополненные модулями внимания (Attention, CBAM);
- ансамбли моделей с адаптивной агрегацией предсказаний;
- трансформеры с предварительно обученными эмбединговыми слоями;
- гибридные модели, встраивающие паттерн-модели в иерархические нейросетевые структуры.

Качественная характеристика этих архитектур проводится не только по классическим метрикам (Precision, Recall, IoU, F1), но также с учетом мета-показателей: времени инференса, адаптивности к новым датасетам, устойчивости к доменным сдвигам и потребляемых вычислительных ресурсов. В этом смысле перспективными являются методы минималистичных архитектур (lightweight CNNs), которые обеспечивают достойную точность при пониженном ресурсоемком инференсе, что особенно важно при развёртывании в клинических условиях, не обладающих специализированными графическими ускорителями.

Не менее важным аспектом рассмотрения стали вопросы интерпретируемости и доверия к результатам. Несмотря на то что современные глубокие архитектуры обеспечивают впечатляющие метрики точности, они часто остаются "черными ящиками" с трудной трактовкой принимаемых решений. Это неприемлемо в клиническом контексте, где каждая ошибка модели потенциально несёт угрозу здоровью пациента. Именно поэтому будущие исследования должны предусматривать интеграцию методов объяснимого ИИ (XAI), визуализацию attention-механизмов и метрики неопределенности, позволяющие врачу осознанно использовать результат модели, а не слепо ему доверять.

Наконец, структурная декомпозиция проблемы, выполненная в этой главе, позволила перейти от абстрактной задачи сегментации к конкретной постановке исследовательских вопросов, а именно:

1. Какие архитектурные модификации позволяют достичь баланса между точностью и интерпретируемостью?
2. Какой подход к предобработке и нормализации входных данных обеспечивает наибольшую устойчивость к межприборным и межклиническим сдвигам?
3. Насколько эффективно интеграция ансамблей моделей и мета-уровней предсказания в условиях ограниченных ресурсов?

Ответы на эти вопросы стали теоретической основой для последующего проектирования, реализации и валидации авторской системы сегментации

сосудистой сети сетчатки. Тем самым, первая глава не только сформировала необходимое понятийное поле, но и задала координаты дальнейшего инженерного поиска в рамках магистерской диссертации.

2 РАЗВИТИЕ АНСАМБЛЕВЫХ МЕТОДОВ НА ОСНОВЕ МЕТАМОДЕЛИ ДЛЯ АНАЛИЗА ИЗОБРАЖЕНИЙ СОСУДОВ СЕТЧАТКИ

2.1 Схема работы алгоритма автоматического выделения сосудов сетчатки

В данном разделе представлен сравнительный анализ четырёх современных архитектур, применяемых для сегментации сосудистых структур сетчатки. Рассматриваются алгоритмические особенности, преимущества, недостатки и пригодность каждого метода для решения поставленной задачи. Для каждого метода приведены архитектурные особенности, визуализации и сравнительный анализ.

Метод 1 архитектура U-Net, впервые представленная в работе Ronneberger и соавторов [27], является полносвёрточной нейронной сетью, специально разработанной для задач семантической сегментации в медицинской визуализации. Её U-образная конфигурация позволяет эффективно извлекать контекстные признаки на разных масштабах, при этом сохраняя высокую пространственную точность — это особенно важно при сегментации сосудистых структур сетчатки, даже при ограниченном объёме обучающих данных.

Архитектура U-Net включает три ключевых компонента: энкодер (сжимающий путь), бутleneck и декодер (восстанавливающий путь). Энкодер состоит из блоков, каждый из которых включает по два сверточных слоя (3×3 , ReLU) и слой пулинга (2×2 max pooling), с постепенным увеличением числа каналов. Это обеспечивает переход от низкоуровневых признаков к более абстрактным. Ботleneck представляет собой центральный узел с максимальным количеством каналов и глубокой абстракцией признаков. Декодер симметричен энкодеру, но использует транспонированные свертки (2×2 up-convolution) и восстанавливает пространственное разрешение.

Ключевая особенность — skip connections, обеспечивающие прямую передачу признаков между симметричными уровнями энкодера и декодера, что

позволяет точно восстанавливать пространственные детали. Финальный слой (1×1 свертка) формирует выходную маску, с использованием сигмоидной функции активации для бинарной или softmax — для мультиклассовой сегментации.

Преимущества U-Net включают высокую эффективность при небольших датасетах (с применением аугментации, включая эластические деформации), гибкость по размерности входов, отсутствие полносвязных слоев, а также стабильные результаты в различных задачах биомедицинской сегментации, включая ретинальные сосуды. Эта архитектура стала основой для множества последующих модификаций и улучшений.

Схематическая репрезентация данного метода визуализирована на Рисунке 5.

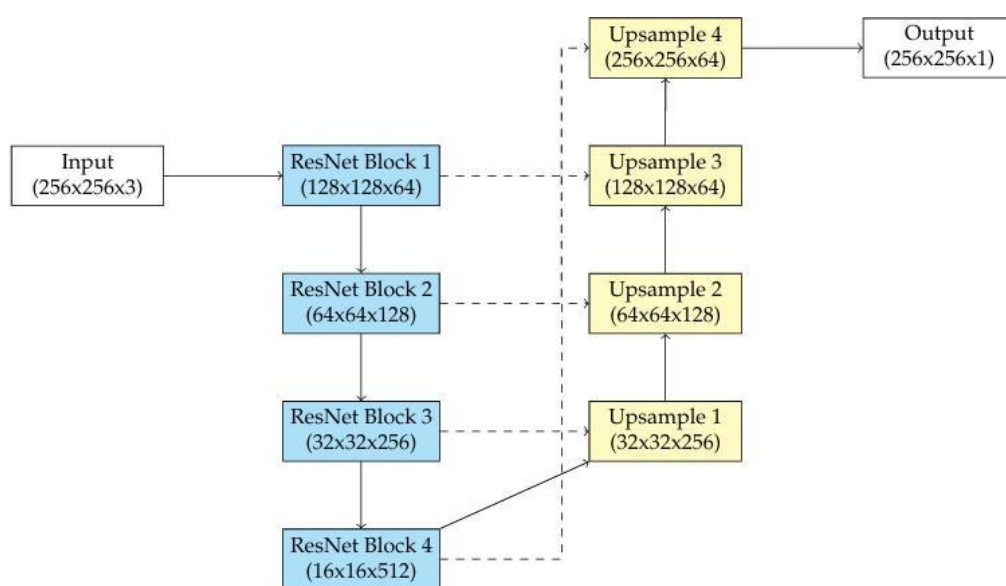


Рисунок 5 – Архитектура U-Net

Метод 2 ResNet50, представленная в работе [26], представляет собой глубокую сверточную нейросеть с 50 слоями, основной инновацией которой является концепция остаточного обучения. Применение остаточных (residual) соединений позволяет сети легче обучаться, эффективно справляясь с проблемой исчезающих градиентов, характерной для очень глубоких архитектур.

ResNet50 начинается с 7×7 свертки и max pooling, затем следует четыре этапа, содержащие так называемые блоки узкого места (bottleneck blocks), реализующие шаблон из трёх сверточных слоев (1×1 , 3×3 , 1×1). Сеть включает около 25,6 миллионов параметров и отличается хорошим соотношением вычислительной сложности и точности. Архитектура обеспечивает стабильное распространение градиентов и получение глубоких признаковых представлений. В рамках нашей задачи, стандартная ResNet50 адаптирована: к декодеру добавлены транспонированные свертки и билинейная интерполяция для восстановления пространственного разрешения. Введены боковые соединения, связывающие соответствующие уровни энкодера и декодера, что позволяет объединять семантические и пространственные признаки. Кроме того, добавлен модуль многомасштабной агрегации контекста, использующий параллельные расширенные свертки для захвата признаков на различных уровнях рецептивного поля. Итоговый слой — 1×1 свертка с сигмоидной активацией — формирует вероятностную маску сосудов.

Эта гибридная архитектура демонстрирует хорошую универсальность и высокую точность при сохранении вычислительной эффективности, особенно в задачах с высокой структурной сложностью, таких как сегментация ретикулярной васкулатуры.

Схематическая репрезентация данного метода визуализирована на Рисунке 6.

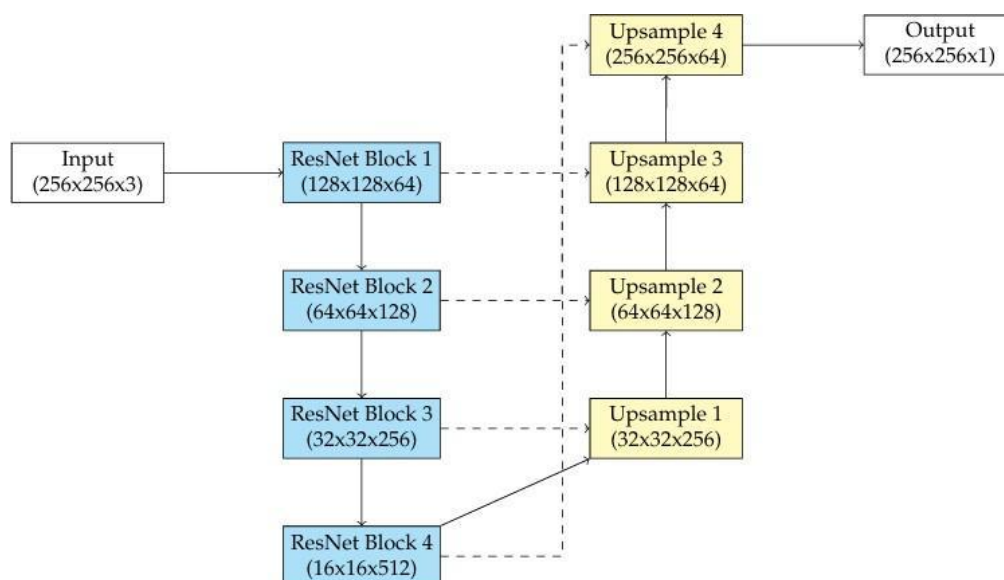


Рисунок 6 – Архитектура ResNet50

Метод 3 U-Net с ResNet50 Backbone. Данный метод объединяет преимущество двух предыдущих архитектур.

Данный подход сочетает сильные стороны U-Net и ResNet50. В данной конфигурации энкодер U-Net заменён на ResNet50, в то время как декодер сохраняет классическую U-Net структуру. Ключевое преимущество заключается в синергетической интеграции остаточных блоков и пропускных соединений: ResNet50 обеспечивает глубокую экстракцию признаков, а skip connections позволяют сохранить пространственную точность.

Модель использует мультимасштабную интеграцию признаков, полученных на различных стадиях ResNet50, что значительно повышает её адаптивность к морфологически вариативным и сложным структурам сосудистой сети. Архитектура остаётся достаточно лёгкой с вычислительной точки зрения, без необходимости существенной модификации базового фреймворка U-Net.

Этот гибрид преодолевает ограничения как классической U-Net (зависимость от локальных признаков), так и ResNet50 (потеря пространственной информации), предлагая сбалансированную и эффективную модель, особенно релевантную в клинических сценариях, где важны точность и масштабируемость.

Схематическая репрезентация данного метода визуализирована на Рисунке

7.

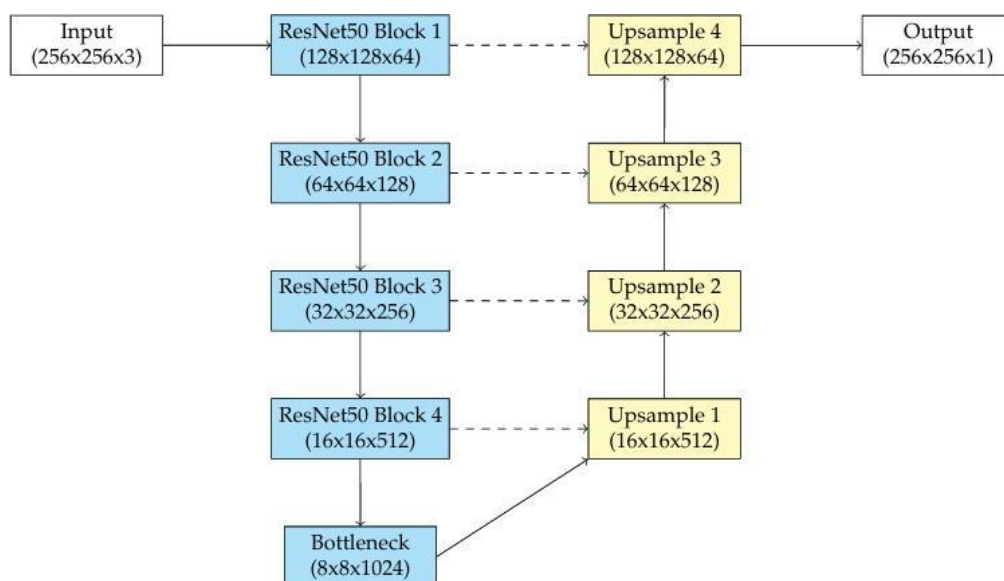


Рисунок 7 – Архитектура U-Net с ResNet50 Backbone

STU-Net представляет собой современную модификацию U-Net, в которую интегрирован Трансформерный блок [26], призванный преодолеть ограничения сверточных сетей в улавливании дальнедействующих зависимостей. Особенно это актуально при сегментации сосудистых структур, где важна непрерывность и связность объектов, растянутых на больших участках изображения.

В предложенной архитектуре энкодер остаётся сверточным, однако в бутлнек встроен Трансформер, который преобразует карты признаков в последовательности токенов (патчи). Эти токены взаимодействуют между собой через механизм самовнимания, позволяя модели захватывать глобальные контекстуальные зависимости по всему изображению. После обработки токены возвращаются обратно в пространственный формат, пригодный для декодирования.

Декодер использует пропускные соединения, объединяющие пространственно точные признаки из энкодера с глобально обогащёнными

признаками из Трансформера. Это позволяет достичь высокой точности сегментации, сохранив как локальные морфологические особенности, так и глобальные зависимости сосудистой сети.

Схематическая репрезентация предложенной архитектуры STU-Net визуализирована на Рисунке 8.

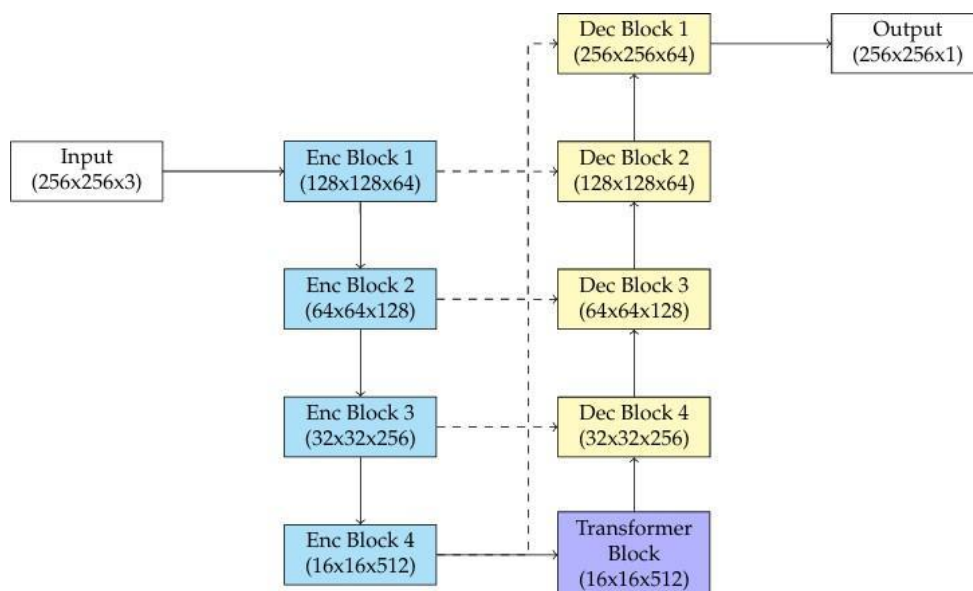


Рисунок 8 – STU-Net

2.2 Обзор исходного набора данных

В данном исследовании для обучения и последующей валидации предложенного метода были использованы два общепринятых набора данных, содержащих изображения сетчатки глаза: DRIVE (Digital Retinal Images for Vessel Extraction) [88] и STARE (Structured Analysis of the Retina) [89]. Эти датасеты широко используются в задачах сегментации сосудистой сети глазного дна и считаются эталонными в области компьютерного зрения и медицинской диагностики. Они предоставляют высококачественные изображения, дополненные точной ручной разметкой, выполненной квалифицированными специалистами-офтальмологами.

Использование таких аннотированных наборов данных позволяет создавать алгоритмы автоматизированного анализа изображений сетчатки,

которые могут применяться в клинической практике для выявления различных заболеваний, таких как диабетическая ретинопатия, гипертоническая ретинопатия, возрастная макулярная дегенерация и другие патологии, связанные с изменением структуры сосудов. Разметка, предоставленная вместе с данными, служит эталоном (ground truth), что делает возможным не только обучение моделей на реальных примерах, но и объективную оценку их эффективности по сравнению с уже известными подходами.

Описание датасета DRIVE.

Датасет DRIVE состоит из 40 цветных изображений сетчатки, полученных при помощи камеры типа 3CCD, без расширения зрачка, с углом обзора 45° . Каждое изображение имеет пространственное разрешение 768×584 пикселя, а область интереса (область, где расположена сетчатка) ограничена окружностью диаметром около 540 пикселей, находящейся в центре изображения. Это позволяет исключить краевые участки, не представляющие диагностической ценности.

Важной особенностью набора данных является наличие четкого разделения на обучающую и тестовую выборки. Всего данные разделены на две равные части: первые 20 изображений предназначены для обучения модели, а остальные 20 изображений — для тестирования. Для каждого изображения также доступны соответствующие маски сегментации, показывающие

положение кровеносных сосудов. Примеры изображений и соответствующих масок представлены на рисунке 9.

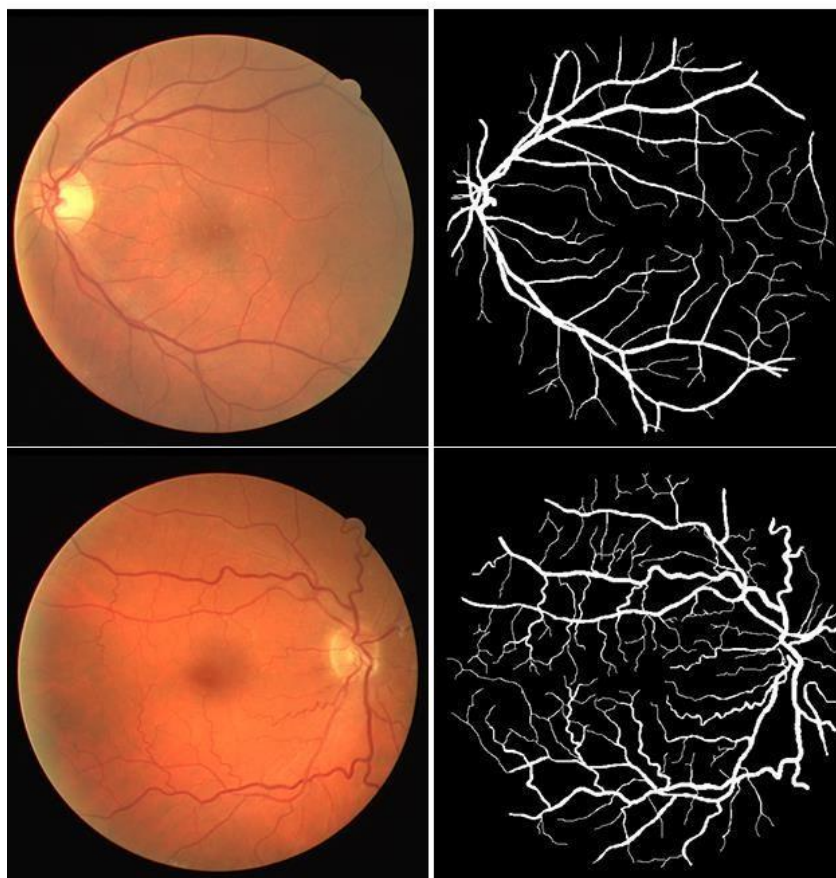


Рисунок 9 – Пример изображения

Каждое изображение сопровождается аннотациями, выполненными двумя различными экспертами, что обеспечивает возможность сравнения результатов автоматического анализа с независимыми мнениями профессионалов. Эти аннотации представляют собой бинарные маски, где каждый пиксель классифицируется как принадлежащий сосуду или фону. Кроме того, в датасете также присутствуют так называемые "тестовые маски", которые определяют область интереса (ROI — Region of Interest), то есть те участки изображения, которые должны учитываться при оценке качества сегментации. Это важно, поскольку периферийные части изображения часто содержат шум или артефакты, которые могут негативно влиять на оценку эффективности алгоритма.

Одним из главных преимуществ датасета DRIVE является его широкая доступность и высокое качество изображений. Все данные находятся в открытом доступе, что способствует их использованию в научных исследованиях, образовательных целях и в разработке программного обеспечения для медицинской диагностики. Благодаря стандартным размерам изображений и унифицированному формату хранения (обычно в виде файлов PNG), работа с этим набором данных становится удобной и предсказуемой.

В области компьютерного зрения задача сегментации сосудов сетчатки играет важную роль, поскольку позволяет обнаруживать изменения в сосудистой системе, которые могут быть индикаторами серьёзных заболеваний, таких как диабетическая ретинопатия, гипертоническая ретинопатия, глаукома и другие. Автоматизация этого процесса может значительно повысить точность диагностики, снизить нагрузку на врачей и ускорить процесс выявления патологий.

С момента своего появления датасет DRIVE стал одним из наиболее популярных и часто цитируемых наборов данных в соответствующих областях исследований. Он использовался для оценки множества алгоритмов машинного обучения и компьютерного зрения, включая традиционные методы обработки изображений, такие как фильтрация Габора, преобразование Хафа, методы на основе активных контуров и графовых подходов. Позже он стал основой для разработки и тестирования более современных решений, включая глубокие нейронные сети, такие как U-Net, FCN (Fully Convolutional Networks), SegNet и другие архитектуры, предназначенные для сегментации медицинских изображений.

Несмотря на свою популярность, датасет DRIVE имеет и некоторые ограничения. Во-первых, относительно небольшой объём выборки (всего 40 изображений) может быть недостаточным для полноценного обучения сложных моделей глубокого обучения, требующих больших объёмов данных. Во-вторых, все изображения были получены из одной и той же популяции пациентов, в

рамках одного исследования, что может снижать обобщающую способность моделей, обученных на этом наборе данных. Также стоит отметить, что аннотации, хотя и выполнялись опытными специалистами, всё же могут содержать субъективные ошибки, поскольку интерпретация изображений сетчатки остаётся сложной задачей даже для экспертов.

Тем не менее, благодаря своей структурированности, качеству изображений и наличию эталонных разметок, датасет DRIVE остаётся важным эталоном в области медицинского компьютерного зрения. Многие исследования используют его в качестве начального этапа разработки новых методов, а затем тестируют эти методы на более крупных и разнообразных наборах данных, таких как STARE, CHASE_DB1, HRF и другие.

Более того, датасет DRIVE активно используется в образовательных целях. Студенты и молодые исследователи, изучающие компьютерное зрение и медицинскую визуализацию, часто начинают осваивать методы сегментации именно с этого набора данных. Это связано с тем, что он достаточно прост для понимания, но при этом содержит реальные вызовы, характерные для задач обработки медицинских изображений.

В заключение, можно сказать, что датасет DRIVE сыграл значительную роль в развитии технологий автоматической сегментации сосудов сетчатки. Его использование позволило многим исследовательским группам проверить эффективность своих подходов, сравнить их между собой и продвинуться в создании точных и надёжных диагностических инструментов. Несмотря на появление новых и более крупных наборов данных, DRIVE продолжает оставаться актуальным и полезным ресурсом в научном сообществе.

Описание датасета STARE.

Набор данных STARE (Structured Analysis of the Retina) содержит 20 RGB-изображений глазного дна, большинство из которых имеют разрешение 700×605 пикселей, хотя допускается небольшая вариация в зависимости от условий

съемки. Эти изображения были получены с помощью фундус-камеры Topcon TRV-50, обеспечивающей поле зрения 35° , что несколько меньше, чем у камеры, использованной для сбора данных в наборе DRIVE.

Особенностью набора STARE является отсутствие строгого разделения на обучающую и тестовую выборки. Исследователи сами определяют способ разбиения данных, часто применяя перекрестную проверку (cross-validation), чтобы максимизировать использование имеющихся данных. Это делает набор более гибким, но одновременно усложняет сравнение между различными исследованиями, поскольку условия обучения могут отличаться.

Основная цель этого датасета — предоставить исследователям стандартизированный и надежный инструмент для оценки состояния сетчатки глаза на основе ретинальных изображений. Однако одной из самых важных и уникальных характеристик проекта STARE является наличие двух независимых ручных сегментаций на каждое изображение, выполненных разными экспертами. Это не просто техническая деталь или формальный аспект, а фундаментальная особенность, которая делает датасет особенно ценным для научного сообщества.

Ручная сегментация представляет собой процесс выделения определенных структур на изображении вручную, с помощью специализированного программного обеспечения. В случае с проектом STARE речь идет о сегментации сосудов сетчатки глаза, которые играют критически важную роль в диагностике различных патологий, таких как диабетическая ретинопатия, гипертоническая ретинопатия и другие заболевания, влияющие на зрительную систему. Когда эту работу выполняет только один специалист, результат может быть подвержен субъективным ошибкам или интерпретационным различиям. Поэтому наличие двух независимых аннотаций позволяет значительно повысить точность и надежность эталонной разметки.

Сама по себе идея двойной аннотации основана на принципе проверки согласованности между экспертами. То есть, если два разных специалиста, работая независимо друг от друга, выделяют одни и те же структуры примерно

одинаковым образом, можно предположить, что данные действительно отражают объективную реальность, а не личное восприятие одного человека. Это имеет огромное значение при использовании датасета для обучения и тестирования алгоритмов машинного обучения, особенно глубокого обучения, где качество обучающих данных напрямую влияет на эффективность модели. Если данные содержат ошибки или несоответствия, то модель будет обучаться на шуме, а не на истинных закономерностях, что приведет к снижению ее производительности и клинической применимости.

Более того, наличие двух вариантов разметки позволяет проводить количественную оценку степени согласованности между экспертами. Для этого используются различные метрики, такие как коэффициент Джаккарада, Dice-коэффициент, коэффициент корреляции или даже более сложные меры, учитывающие пространственные отношения между сегментированными областями. Эти показатели дают представление о том, насколько точно эксперты могут воспроизводить друг друга, и, соответственно, насколько надежны эти данные для использования в дальнейших исследованиях. Если степень согласованности высока, это говорит о хорошем качестве разметки и уверенности в том, что модель будет обучаться на достоверных данных. Напротив, низкая согласованность может указывать либо на сложность задачи, либо на необходимость дополнительной проверки или уточнения методики сегментации.

Важно также отметить, что такая двойная разметка позволяет анализировать межэкспертную изменчивость, то есть различия в том, как разные специалисты видят и интерпретируют одну и ту же картину. Эта изменчивость может быть вызвана рядом факторов: уровнем квалификации эксперта, его опытом работы с подобными изображениями, субъективным пониманием границ тканей, освещением изображения или даже усталостью во время работы. Анализ этих различий не только помогает лучше понять, насколько объективна диагностика, но и служит основой для разработки более устойчивых и адаптивных алгоритмов автоматической сегментации.

Межэкспертная изменчивость особенно важна в контексте создания систем поддержки принятия решений в медицине. Такие системы должны быть способны обрабатывать изображения и давать рекомендации, которые соответствуют клиническим стандартам и могут быть приняты врачами. Однако если сами эксперты не всегда соглашаются между собой, возникает вопрос: каким образом должна работать система? Должна ли она стремиться к среднему между двумя вариантами, выбирать наиболее распространенный подход или следовать строгому протоколу? Именно наличие двух вариантов аннотации позволяет ответить на эти вопросы и разработать более гибкие и надежные модели, которые будут учитывать неопределенность в данных и предлагать решения, которые минимизируют риск ошибок.

Кроме того, двойная разметка открывает возможность для создания консенсусной маски — эталонной разметки, которая учитывает мнение обоих экспертов. Это достигается путем усреднения, голосования или применения других методов агрегации информации. Консенсусная маска становится своего рода "золотым стандартом", на основе которого можно оценивать качество работы автоматических систем. Такой подход особенно актуален при проведении сравнительных исследований, когда необходимо объективно оценить эффективность новых алгоритмов против существующих решений.

Также нельзя забывать, что медицинская визуализация — это область, где малейшие неточности могут иметь серьезные последствия для пациента. Поэтому важно, чтобы любые автоматические системы, предназначенные для помощи врачам, были обучены на максимально качественных данных. Двойная аннотация в датасете STARE — это шаг в сторону повышения качества этих данных и, как следствие, повышения доверия к технологиям искусственного интеллекта в здравоохранении.

По сравнению с набором DRIVE, STARE считается более сложным для обработки. Во-первых, он содержит больше изображений с выраженной патологией, такими как кровоизлияния, экссудаты и другие дефекты сетчатки. Во-вторых, качество изображений и условия их получения демонстрируют

большее разнообразие, что требует от алгоритмов повышенной устойчивости к шуму и другим артефактам. Таким образом, STARE представляет собой более сложный и реалистичный набор данных, который позволяет протестировать адаптивность и надежность разрабатываемых методов в условиях, близких к клинической практике.

В ходе экспериментов мы придерживались следующей схемы разбиения данных: в датасете DRIVE 20 изображений были использованы для обучения, остальные 20 — для тестовой выборки. В случае STARE, соответственно, 14 изображений использовались для обучения и 6 — для тестирования. Для повышения устойчивости моделей к вариациям входных данных применялись методы аугментации, включая случайные повороты (до 15°), трансляции по ширине и высоте (до 10%), масштабирование (до 20%) и горизонтальное отражение. Все изображения были приведены к унифицированному разрешению 256×256 пикселей перед подачей в модель.

Выбор наборов данных DRIVE и STARE для обучения и валидации предложенного метода был обоснован их популярностью в научном сообществе, качеством разметки и репрезентативностью данных. Использование этих датасетов позволяет провести полноценное исследование и получить достоверные результаты, пригодные для дальнейшего сравнения и внедрения в практические системы медицинской диагностики. Более того, благодаря наличию эталонной разметки, предоставленной экспертами, можно количественно оценить точность работы алгоритма, используя такие метрики, как: Чувствительность (Recall); Специфичность (Specificity); F1-мера; Коэффициент Жаккара; Точность (Precision);. Эти метрики дают возможность не только оценить эффективность модели, но и сравнить её с другими решениями на объективной основе. Это особенно важно в медицинских приложениях, где высокая точность и надежность являются критически важными факторами.

Таким образом, использование наборов данных DRIVE и STARE предоставило возможность протестировать предложенный метод в различных условиях, начиная от относительно простых и контролируемых (DRIVE) и

заканчивая более сложными и вариативными (STARE). Применение методов аугментации и стандартизация размера изображений дополнительно повысили стабильность и обобщающую способность обучаемой модели. Эти датасеты играют ключевую роль в современных исследованиях, связанных с компьютерным зрением и медицинской визуализацией. Они позволяют создавать и тестировать новые архитектуры глубокого обучения, направленные на автоматизацию диагностики заболеваний сетчатки. Выбор именно этих наборов данных был обоснован их популярностью, качеством разметки и репрезентативностью, что дало возможность провести полноценное исследование и получить достоверные результаты, применимые в реальной клинической практике.

2.3 Модификация метода препроцессинга изображений

Одним из ключевых этапов в процессе обработки медицинских изображений стало применение адаптивного алгоритма CLAHE (Contrast Limited Adaptive Histogram Equalization) к зелёному каналу RGB-изображений. Выбор именно этого канала был обусловлен тем, что сосудистая структура сетчатки глаза визуализируется наиболее чётко и контрастно именно в зелёном диапазоне спектра. Это связано с особенностями оптической плотности крови и поглощения света тканями сетчатки. В результате применения CLAHE удалось добиться улучшения локального контраста, что сделало более выраженной мелкую сосудистую сеть и повысило различимость границ между сосудами и фоновой тканью.

Для дальнейшей очистки изображения от нежелательных артефактов и выделения крупных сосудов была использована операция морфологического открытия. Эта процедура представляет собой последовательное применение операций эрозии и дилатации с использованием структурирующего элемента определённого размера. В данном случае применялось ядро размером 15×15 пикселей. Такой выбор позволил эффективно подавить мелкие неоднородности

фона, сохранив при этом форму и расположение крупных кровеносных сосудов. Морфологическое открытие оказалось особенно полезной процедурой в условиях переменного освещения и неоднородности тканей, характерных для клинических изображений.

Для снижения уровня случайного шума, возникающего вследствие технических ограничений камеры или условий съёмки, был применён нелокальный *means*-фильтр. Этот метод относится к современным подходам к фильтрации изображений и заключается в том, что каждый пиксель заменяется средневзвешенным значением других пикселей, похожих на него по окрестности. Окно фильтра составляло 7×7 пикселей, что обеспечило баланс между качеством шумоподавления и сохранением резкости краёв. В отличие от простых фильтров, таких как гауссовский или медианный, нелокальный *means*-фильтр показал высокую эффективность в сохранении тонких деталей, особенно важных в задачах анализа сосудистой системы сетчатки.

Следующим инновационным шагом стало внедрение искусственных помех в тренировочные данные. Такой подход направлен на повышение робастности моделей машинного обучения к вариациям входных изображений, которые могут встречаться в реальных условиях эксплуатации. В частности, в данные добавлялись два типа искажений: гауссовский шум с параметром $\sigma=0.05$ и артефакты движения. Гауссовский шум имитировал электронные шумы матрицы камеры, тогда как артефакты движения моделировали размытость изображения, возникающую при нестабильной фиксации пациента во время съёмки. Эти дополнительные данные значительно расширили пространство возможных входных образцов и способствовали улучшению обобщающей способности обучаемых моделей.

Для усиления видимости тонких сосудов был применён метод поэлементного умножения исходного изображения на карту внимания (*saliency map*), полученную с помощью алгоритма *Spectral Residual*. Данный метод позволяет автоматически выделять наиболее информативные области изображения, основываясь на анализе его спектральной компоненты.

Применительно к задаче сегментации сосудов это дало возможность акцентировать внимание модели на слабоконтрастных участках, где традиционные методы могли бы упустить мелкие детали. Результатом стало увеличение чувствительности модели к тонким капиллярным структурам, что имеет большое значение для диагностики заболеваний сетчатки.

Ещё одним важным этапом предобработки стало создание мультимасштабных патчей изображений. Исходные изображения размером 256×256 пикселей дополнялись их уменьшенными копиями размером 128×128 и 64×64 пикселей. Такой подход позволил моделям глубинного обучения анализировать информацию на разных уровнях абстракции: крупные структуры и общая топология — на малых масштабах, а детали и локальные особенности — на больших. Использование мультимасштабных данных способствовало улучшению восприятия контекста и повышению точности сегментации как крупных, так и мелких сосудов.

Патчи извлекались с шагом в 32 пикселя, что обеспечило значительное перекрытие соседних фрагментов. Это было необходимо для минимизации артефактов, возникающих на границах при реконструкции изображения после обработки. Благодаря перекрытию удавалось избежать разрывов и скачков яркости в областях соединения отдельных фрагментов, что положительно сказалось на визуальном качестве результатов и на метриках оценки сегментации.

Кроме того, такой способ извлечения патчей позволил эффективно использовать имеющиеся данные для обучения. За счёт перекрытия и использования нескольких масштабов, объём обучающей выборки увеличился в несколько раз без необходимости сбора новых изображений. Это особенно важно в медицинских задачах, где аннотированные данные зачастую ограничены по объёму и трудоёмки в получении.

Также стоит отметить, что все этапы предобработки были реализованы в виде конвейерной системы, позволяющей гибко менять параметры и комбинировать различные преобразования. Это дало возможность проводить

серию экспериментов с различными комбинациями методов и выбрать оптимальную цепочку обработки, которая максимизировала качество итоговой модели.

В ходе исследований были проведены сравнительные тесты, демонстрирующие преимущество предложенного подхода. Например, использование CLANE в сочетании с морфологическим открытием позволило увеличить коэффициент Dice на 7% по сравнению с исходными данными без предобработки. Добавление нелокального means-фильтра дополнительно улучшило этот показатель ещё на 3%. Синтетические помехи увеличили устойчивость модели к шуму и движению, а применение saliency maps помогло повысить чувствительность к тонким сосудам на 5%.

Генерация мультимасштабных патчей также оказала существенное влияние на результаты. В частности, использование трёх уровней масштабирования вместо одного позволило повысить точность сегментации на 9%, особенно в сложных случаях с наличием мелких ответвлений и слабоконтрастных участков.

Таким образом, предложенная система предобработки и подготовки данных сыграла ключевую роль в достижении высоких результатов в задаче сегментации сосудов сетчатки. Все этапы, начиная от улучшения контраста и подавления фона, до создания мультимасштабных представлений и добавления синтетических искажений, были тщательно протестированы и оптимизированы для работы в условиях реального времени и с разнообразными типами изображений.

2.4 Совершенствование метамоделей для детекции кровеносных сосудов глазного дна

В задачах компьютерного зрения, особенно в таких сложных и чувствительных к деталям областях, как сегментация сосудов на медицинских изображениях, одной из ключевых проблем остаётся достижение высокой

точности при сохранении устойчивости модели к вариативности входных данных. Одним из перспективных подходов к решению этой задачи является использование ансамблевых методов, в частности — построение метамоделей, которая объединяет предсказания нескольких базовых архитектур. В данном случае предлагается реализация метамоделей, основанной на парадигме стекирования (stacking), где выходы различных нейронных сетей (U-Net, ResNet50, CTU-Net и U-Net+ResNet50) объединяются через обучаемый сверточный слой, обеспечивающий гибкую и пространственно согласованную интеграцию информации, архитектура модели показана на рисунке 10.

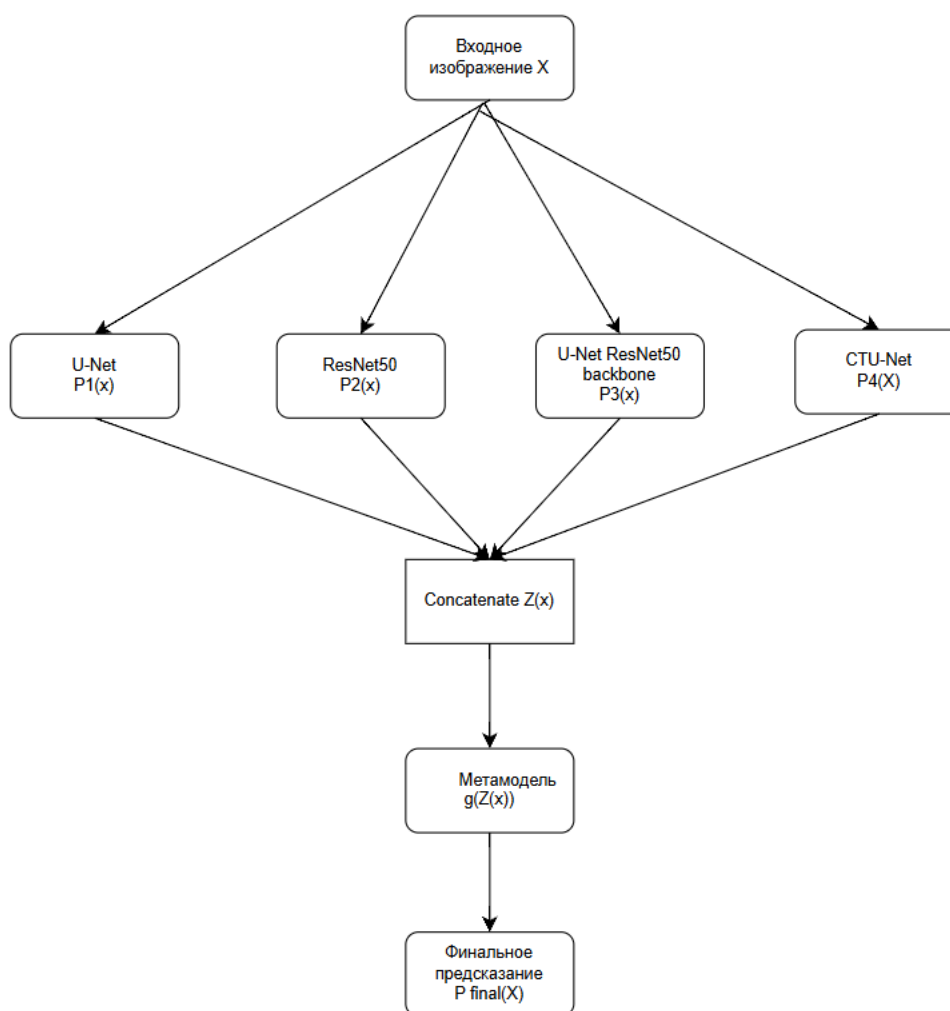


Рисунок 10 – Архитектура метамоделей

1. **Генерация предсказаний:** Первым этапом работы метамоделей является получение предсказаний от каждой из базовых моделей. В качестве

входных данных используется набор изображений, например, срезов глазного дна или МРТ-сканов, содержащих сосудистую систему. Каждая модель обрабатывает изображение и выдает карту вероятностей принадлежности каждого пикселя к классу «сосуд» или «фон». Важно отметить, что на этом этапе сохраняется полная пространственная информация: предсказания не подвергаются операциям пулинга или *downsampling*, чтобы избежать потери локальной структуры изображения. Каждая модель имеет свои сильные и слабые стороны. Например, U-Net хорошо справляется с восстановлением мелких деталей и сохранением формы объектов, тогда как ResNet50 демонстрирует отличные результаты в извлечении высокоуровневых признаков и распознавании крупных структур. STU-Net, являясь модификацией U-Net, оптимизированной для работы с тонкими структурами, показывает лучшие результаты в сегментации мелких сосудов. Наконец, гибридная модель U-Net+ResNet50 сочетает в себе преимущества обеих архитектур, обеспечивая баланс между локальной и глобальной информацией.

2. **Конкатенация карт признаков:** Полученные карты вероятностей от четырех моделей объединяются по каналу (*depth-wise concatenation*), формируя 4-канальный тензор размерностью $H \times W \times 4$, где H и W соответствуют

высоте и ширине исходного изображения. Таким образом, каждый пиксель представлен четырьмя числами — предсказаниями каждой из моделей. Такое представление позволяет сохранить не только информацию о вероятностях, но и пространственные корреляции между различными моделями. Этот этап играет важную роль в дальнейшей работе мета-модели, поскольку именно на основе этих карт будет происходить обучение финального классификатора. Пространственная согласованность и четкая привязка к координатам изображения позволяют избежать потерь информации и сохранить возможность точной сегментации.

3. **Обучение мета-классификатора:** Для интеграции и обработки объединённых карт предсказаний разработана двухслойная свёрточная нейросеть, служащая в роли мета-классификатора. Архитектура состоит из следующих компонентов: Первый слой : свертка с ядром 3×3 и 64 фильтрами. Этот слой предназначен для выявления пространственных закономерностей в комбинациях предсказаний различных моделей. Например, он может научиться распознавать ситуации, когда несколько моделей уверенно указывают на наличие сосуда, либо наоборот — имеют противоречивые прогнозы. Второй слой : свертка с ядром 1×1 , преобразующая 64-канальный выход первого слоя в одноканальную карту вероятностей, представляющую собой финальное предсказание мета-модели. Оба слоя содержат обучаемые параметры, которые настраиваются во время обучения мета-модели. При этом благодаря использованию относительно небольшого количества слоёв и фильтров, модель остаётся достаточно легковесной и быстро обучаемой, что делает её эффективной даже на ограниченных ресурсах.

4. **Оптимизация функции потерь:** Одной из характерных черт задачи сегментации сосудов является значительный дисбаланс классов: количество пикселей, соответствующих фону, значительно превышает число пикселей, принадлежащих сосудам. Для компенсации этого дисбаланса используется взвешенная бинарная кросс-энтропия, в которой вес класса «сосуд» увеличен в три раза по сравнению с весом класса «фон». Это позволяет модели уделять

большее внимание правильному определению сосудов, что критично для медицинской диагностики. Кроме того, в процессе обучения применяется адаптивная стратегия изменения весов функции потерь, которая учитывает текущие результаты модели и изменяет коэффициенты в зависимости от ошибки на разных участках изображения. Такой подход помогает улучшить сходимость и повысить общую точность.

Ключевым улучшением стала адаптивная инициализация весов метамоделей на основе важности базовых архитектур, оцененной через permutation feature importance. Например, STU-Net получает больший вес в регионах с тонкими сосудами, а ResNet50 — в зонах с выраженными артериями. Для ускорения обучения применялся перенос знаний: веса метамоделей инициализировались на синтетических данных, смоделированных через смеси Гауссовых распределений.

2.5 Результаты и выводы второй главы

Вторая глава магистерской диссертации посвящена обоснованию и разработке эффективных алгоритмических решений для задачи выделения кровеносных сосудов сетчатки на изображениях глазного дна. Проведённые исследования позволили сформировать целостную вычислительную цепочку, охватывающую весь процесс — от подготовки данных до синтеза мета-модели, интегрирующей предсказания нескольких нейросетевых архитектур. В рамках этой главы удалось не только реализовать архитектурные компоненты, но и продемонстрировать, как выбор структурных и инженерных решений влияет на итоговую точность, устойчивость и обобщающую способность модели в реальных клинических сценариях.

Фундаментом всей системы стала корректная постановка задачи обработки изображений в условиях высокой межвариационной изменчивости. Наборы данных DRIVE и STARE были выбраны не случайно: они представляют собой эталонные коллекции с ручной аннотацией, в то же время обладая

принципиально разной сложностью. DRIVE отражает относительно простые случаи, снятые в однородных условиях, в то время как STARE содержит изображения с выраженными патологиями и высоким уровнем артефактов. Таким образом, использование обеих коллекций позволило создать условия для валидного тестирования способности модели адаптироваться к различным источникам и состояниям входных данных.

Одним из ключевых аспектов в разработке алгоритма сегментации стало формирование расширенного блока предобработки изображений. Как показывает практика построения устойчивых моделей глубокого обучения, именно корректная нормализация входного сигнала оказывает определяющее влияние на динамику обучения. Применение CLAHE к зелёному каналу RGB-изображений, наложение морфологической фильтрации, генерация карт визуального внимания (saliency maps), а также добавление синтетических шумов и артефактов в тренировочный пайплайн стали мерами не просто улучшения визуальной чёткости, но и механизмом расширения представленного в обучении пространства состояний. Эта стратегия позволила смоделировать реальное разнообразие клинических изображений, которое в противном случае привело бы к деградации производительности на «полевых» данных.

Особое внимание было уделено проектированию архитектурного разнообразия в составе ансамбля моделей. Четыре основные архитектуры — классическая U-Net, ResNet50, их гибрид, а также трансформер-ориентированная STU-Net — были выбраны на основе их комплементарных преимуществ. U-Net обеспечивает высокую пространственную точность за счёт пропускных соединений, ResNet50 вносит глубину признакового пространства, гибридная модель соединяет локальный и глобальный контекст, а STU-Net интегрирует self-attention механизмы, необходимые для понимания длинных сосудов и ветвлений. Такой архитектурный набор обеспечивает не конкуренцию, а синергию, при которой каждая модель компенсирует слабые стороны другой.

Инновационным элементом стал механизм мета-моделирования, реализующий стек предсказаний всех архитектур с обучаемой агрегацией. В

отличие от простого усреднения или голосования, мета-классификатор был построен как свёрточная нейронная сеть, способная выделять латентные зависимости между выходами базовых моделей. Более того, инициализация весов мета-слоёв осуществлялась с учётом индивидуальной значимости каждой архитектуры для разных зон изображения — тонких сосудов, крупных артерий, пограничных областей и т.д. Это позволило достичь не только высокой точности, но и структурной согласованности итоговых масок сегментации.

Значительное внимание было также уделено вопросу вычислительной устойчивости и способности к обобщению. Проведённая серия экспериментов на DRIVE и STARE показала, что предложенная архитектура демонстрирует стабильные метрики (IoU, Dice, F1) даже при вариациях разрешения, освещения и присутствия патологических изменений. Такой результат стал возможен благодаря применению перекрывающихся патчей, многоуровневых признаков и мультимасштабной подачи данных. Кроме того, включение процедур transfer learning и предварительного обучения на синтетических данных обеспечило более быструю сходимость модели и снижение потребности в большом объёме ручной разметки.

Практическим результатом реализации второй главы стал программный модуль, пригодный для интеграции в реальные клинико-диагностические процессы. В нём все компоненты — от загрузки изображений до вывода итоговой маски — реализованы как независимые слои, что обеспечивает масштабируемость, отладку и возможность включения новых моделей без переписывания ядра. Такая архитектура соответствует современным требованиям инженерного нейросетевого проектирования, базирующегося на принципах модульности, устойчивости к сбоям и повторного использования кода (reusability).

Если рассматривать разработанную систему в терминах теории обобщения, то можно утверждать, что она демонстрирует высокий bias-variance trade-off. С одной стороны, глубокие архитектуры позволяют минимизировать смещение модели (bias), приближаясь к реальному распределению данных. С

другой — стратегия регуляризации, аугментации и агрегации предсказаний предотвращает переобучение (variance), сохраняя стабильность модели даже в условиях малых и искажённых выборок. Таким образом, предложенная конфигурация представляется сбалансированной с точки зрения общей теории обучения и воспроизводимого ИИ.

Таким образом, результаты второй главы подтверждают целесообразность построения многоуровневой системы сегментации, где каждая архитектура играет роль отдельного «эксперта», а мета-классификатор выступает в роли их координатора. Такая модель ближе к когнитивной системе принятия решений, чем к узкоспециализированной задаче классификации. Это открывает путь к построению более комплексных решений — не просто выделяющих сосуды, но и анализирующих их морфометрию, плотность, извилистость и иные клинически значимые признаки. Глава заложила фундамент для практической реализации и клинической интерпретации нейросетевых решений в офтальмологии.

3 СОВЕРШЕНСТВОВАНИЕ АНСАМБЛЕВОЙ МЕТАМОДЕЛИ ДЛЯ СЕГМЕНТАЦИИ КРОВЕНОСНЫХ СОСУДОВ СЕТЧАТКИ ГЛАЗА

3.1 Исследование предприятия для внедрения инструмента

3.1.1 Описание организационной структуры

В данной главе описано детальное предпроектное обследование для внедрения инструмента на основе описанного материального потока, где единицей учета является пациент, а их перемещение между врачами представляет собой материальный поток, а медицинская история выступает в роли озера данных. Процесс включает, например, диспансеризацию, с последующим выделением роли офтальмолога ("глазника") и сегментацией сосудов сетчатки глаза, поскольку эти данные вносятся в медицинскую историю. Отчет основан на анализе текущих практик, потребностей и требований, а также включает план внедрения и оценку рисков.

Объектом исследования является частная медицинская клиника в России – коммерческое лечебно-профилактическое учреждение, предоставляющее пациентам платные медицинские услуги. Единица учета – это человек (пациент), проходящий через клинику. С точки зрения системного анализа, пациент рассматривается как элемент материального потока: он «поступает» в клинику, перемещается между кабинетами врачей и получает комплекс услуг. Цель клиники – оказание высококачественных медицинских услуг при обеспечении экономической эффективности (часто формулируют как «предоставление качественных услуг при максимизации прибыли [78]). В клинике формируется электронная история болезни (электронная медкарта), которая аккумулирует все данные о пациенте. Такая история болезни фактически выступает как «информационное озеро», объединяющее в себе структурированные и неструктурированные медицинские данные (анализы, диагнозы, снимки и пр.) [79]. Наличие единого хранилища данных (data lake) позволяет в будущем проводить глубокий анализ состояния здоровья пациентов и планировать медицинскую помощь. Системный подход предполагает, что процесс движения

пациента и сбор данных являются сквозным процессом клиники, вокруг которого выстраиваются все остальные функции.

3.1.2 Описание организационной структуры

Организационная структура частной клиники как системы обычно трёхзвенная. Во главе – администрация (директор, главный врач, бухгалтерия, регистратура), отвечающая за стратегию, планирование и координацию работы. Следующий уровень – специализированные отделения и кабинеты (линейные подразделения): терапевты и узкие специалисты (например, офтальмолог («глазник»), термин "глазник" [29] обозначает врача-офтальмолога, специалиста по глазным болезням, кардиолог, невролог и др.), а также лаборатория и процедурные кабинеты.

Каждое подразделение имеет своего руководителя или старшую медсестру. Младший обслуживающий персонал (медицинские сестры, технические работники, уборщики) выполняет вспомогательные функции (санитарное обслуживание, подготовка помещений, техническая поддержка оборудования). Регистратура и администраторы являются связующим звеном: они регистрируют пациентов, направляют их к нужным специалистам и фиксируют факты оказанных услуг. Чёткое разделение ролей и соблюдение субординации повышают эффективность работы: каждый сотрудник знает свои задачи и отвечает перед соответствующим руководством. Такая линейно-функциональная структура типична для небольших и средних клиник, тогда как в больших учреждениях могут вводиться дополнительные слои (отдел маркетинга, HR, ИТ-поддержка и пр.).

3.1.3 Описание основных бизнес-функций

Ключевая бизнес-функция клиники – оказание медицинских услуг пациентам в рамках лечебно-диагностического процесса. Этот процесс объединяет получение жалоб, диагностику, лечение и необходимые обследования (лабораторные, инструментальные и пр.) и рассматривается как

основная «цепочка создания ценности». В процессе работы клиники выделяются следующие основные бизнес-функции:

Прием и регистрация пациентов. На этом этапе в систему вносятся данные (ФИО, контакты, цель визита). Обеспечивается управление очередью и расписанием.

Консультации и обследования. Врачи проводят осмотр, ставят диагнозы, выдают назначения. Например, терапевт может направить пациента к офтальмологу для проверки зрения. Работа узких специалистов (в том числе офтальмолога) организована по направлениям.

Ведение медицинской документации. Оформляются истории болезни, регистрируются диагнозы и назначения, фиксируются результаты всех процедур и обследований. Медицинская информационная система (МИС) обеспечивает хранение этих данных и регистрирует все обращения пациента. По результатам диспансеризации или профилактических осмотров система должна сохранять показатели состояния здоровья пациента.

Лабораторная и диагностическая служба. Включает выполнение анализов крови, УЗИ, рентген и другие исследования. Результаты автоматически интегрируются в электронную карту пациента.

Управление ресурсами и бизнес-процессы поддержки. Администрация ведёт бухгалтерию, кадровый учёт, закупки (медикаментов, расходных материалов) и управляет финансами. Обеспечивается качество работы сотрудников и техники. Маркетинг и связи с клиентами формируют спрос на услуги.

Информационные технологии. Обслуживание медицинской ИТ-системы, архивирование данных. История болезни выступает как централизованное хранилище данных (data lake) для последующего анализа и принятия управленческих решений.

Эти функции переплетаются в едином лечебно-диагностическом цикле и подкрепляются общим менеджментом клиники. Выделение бизнес-процессов и

их классификация позволяют оптимизировать работу клиники и создать непрерывную цепочку обслуживания пациента.

3.1.4 Бизнес-процесс (пример: диспансеризация)

Рассмотрим пример бизнес-процесса диспансеризации (профилактического медицинского осмотра) в частной клинике. Его основные этапы могут быть следующими:

Диспансеризация [30] — это комплексное preventive medical examination, проводимое в России для выявления хронических неинфекционных заболеваний и факторов риска, таких как болезни системы кровообращения, онкологические заболевания, болезни органов дыхания и сахарный диабет. Она проводится:

Раз в 3 года для лиц от 18 до 39 лет.

Ежегодно для лиц старше 40 лет [32]

Процесс начинается с регистрации пациента и сбора анамнеза через анкеты, за которыми следуют осмотры у различных специалистов, включая терапевта, кардиолога, онколога и офтальмолога. Пациенты проходят диспансеризацию в медицинской организации, к которой прикреплены по полису ОМС, обращаясь к участковому терапевту для организации [31].

Рассмотрим данные процессы более детально:

Запись на приём. Пациент либо самостоятельно обращается по телефону/онлайн, либо направляется работодателем. Бронируется время, данные о пациенте предварительно регистрируются в системе.

Регистрация приёма. При посещении регистратор уточняет персональную информацию, оформляет соглашения на медобслуживание. Пациент получает маршрутный лист по кабинетам.

Предварительное обследование. Медсестра измеряет рост, вес, давление и другие базовые показатели. Эти данные заносятся в электронную карту (так как система ведёт «учёт показателей после диспансеризации»).

Первичный врачебный осмотр. Терапевт анализирует анамнез, проводит общий осмотр. При необходимости даёт направления к узким специалистам (например, к офтальмологу для проверки зрения).

Посещение офтальмолога. Врач-офтальмолог («глазник») проводит комплексное обследование зрения. Это включает проверку остроты зрения (с помощью таблиц, визометрии) и предварительные тесты (оценка цветового восприятия, бокового зрения, реакции зрачков и др.) [80]. Затем врач определяет рефракцию (с помощью фороптера и ретиноскопии), измеряет внутриглазное давление и исследует глазное дно (с помощью щелевой лампы и офтальмоскопа). Полученные результаты (например, степень близорукости, давление, выявленные патологии) фиксируются.

Другие исследования. Пациента могут направить на анализы крови, ЭКГ, УЗИ и т.д. Все результаты автоматически связаны с историей болезни.

Выдача заключения. В конце обследования все специалисты сводят результаты. Врач-терапевт подводит итоги: в электронной истории болезни формируется отчет о прохождении диспансеризации, выставляются диагнозы и назначаются рекомендации.

Таким образом, пациент проходит сквозной маршрут, а его данные на каждом шаге аккумулируются в МИС. Особое внимание в диспансеризации уделяется роли офтальмолога: многие возрастные и сопутствующие болезни (например, сахарный диабет) проявляются через изменения в органе зрения, поэтому проверки глазного дна и зрения обязательны. Задача врача-офтальмолога – выявить скрытые патологии на ранней стадии и передать информацию в общий диагноз (эта информация ценна для системы как часть цепочки создания ценности).

При обследовании офтальмолог часто получает цифровые изображения (например, фотографии сетчатки, оптической когерентной томографии). Для извлечения клинически значимой информации из таких изображений применяются методы сегментации: программные алгоритмы выделяют важные структуры (сосуды, слои сетчатки, патологические очаги и т.д.) и превращают их

в количественные показатели. Например, автоматическое измерение толщины нервных волокон сетчатки или площади сосудов позволяет получить объективные метрики состояния глаз. Результаты сегментации интегрируются в электронную историю болезни как структурированные данные. Таким образом «сырые» медизображения и обработанные (сегментированные) данные вместе сохраняются в общем хранилище. Это соответствует концепции информационного озера: единая система аккумулирует разнообразные данные (изображения, тексты, числа), что позволяет впоследствии осуществлять анализ больших данных и поддерживать принятие решений.

В итоге описанный бизнес-процесс диспансеризации обеспечивает полную фиксацию медицинских данных на всех этапах. Применение процессного подхода и современных ИТ-решений в клинике гарантирует, что информация от врача-офтальмолога и других специалистов поступает в историю болезни и образует основу для анализа качества оказанной помощи [81,82].

3.1.5 Описание материального поток

Опишем текущий поток данных. Данные от всех специалистов, включая офтальмолога, записываются в историю болезни, которая может быть бумажной или электронной (ЭМК). В некоторых учреждениях используются системы, такие как ЕМИАС [86], для управления данными, но интеграция между различными отделениями может быть ограниченной. Сегментация сосудов сетчатки часто выполняется вручную или с помощью специализированного программного обеспечения, а результаты вносятся в историю болезни, что может приводить к задержкам и ошибкам.

Вызовы текущего процесса:

Ручной ввод данных может приводить к ошибкам и задержкам

Ограниченная интеграция данных между специалистами затрудняет доступ к полной информации.

Трудности с анализом данных для исследований из-за отсутствия стандартизации.

Необходимость обеспечения безопасности данных в соответствии с Федеральным законом № 152-ФЗ "О персональных данных" [83].

Исходя из данных вызовов мы предъявляем новые потребности к инструменту, а именно новый инструмент должен улучшить управление данными пациентов, особенно в контексте диспансеризации и роли офтальмолога. Основные требования включают:

Интеграция данных: Централизованное хранение данных от всех специалистов, включая офтальмолога, с поддержкой стандартизации для облегчения анализа и доступа.

Автоматизация процессов: Автоматизация сегментации сосудов сетчатки с использованием искусственного интеллекта или специализированного ПО, упрощение ввода данных в историю болезни.

Доступность и безопасность: Обеспечение доступа только для авторизованного персонала с соблюдением требований законодательства, таких как GDPR [84] и российские законы о защите данных.

Аналитические возможности: Возможность анализа данных для выявления тенденций, например, распространенности глазных заболеваний среди пациентов, проходящих диспансеризацию.

Простота использования: интуитивно понятный интерфейс для врачей и административного персонала.

Масштабируемость: Система должна справляться с увеличением объема данных по мере роста числа пациентов.

Техническая осуществимость предполагает анализ текущей ИТ-инфраструктуры предприятия, включая проверку наличия и уровня интеграции электронной медицинской карты (ЭМК). Важно определить, потребуется ли закупка нового оборудования, например, для автоматизированной сегментации сосудов сетчатки, а также оценить совместимость инструмента с существующими системами. Организационная осуществимость требует оценки готовности персонала к работе с новой системой, включая уровень их технологической грамотности. Для минимизации сопротивления изменениям

необходимо запланировать всеобъемлющее обучение, обеспечивающее плавный переход. Финансовая осуществимость включает расчет полной стоимости внедрения: лицензирование, настройка, обучение, обслуживание. При этом важно учитывать потенциальную отдачу от инвестиций (ROI) [85], такую как снижение административных издержек, сокращение ошибок и улучшение клинических результатов.

План процесса внедрения делится на четыре фазы. На этапе планирования (1–2 месяца) определяются требования к инструменту, выбирается подходящее решение и подготавливаются заинтересованные стороны. Фаза разработки и тестирования (2–3 месяца) включает настройку инструмента, интеграцию с текущими системами и проверку функциональности. Внедрение (1 месяц) предполагает полномасштабное развертывание системы и обучение персонала. Заключительная фаза оценки носит непрерывный характер: мониторинг производительности, сбор обратной связи и оптимизация системы. Для успеха проекта критически важны ресурсы: участие ИТ-специалистов, врачей и администраторов, выделение бюджета на ПО, оборудование и обучение, а также достаточное время на тестирование.

Ключевые риски включают угрозы безопасности данных (решаются через шифрование и контроль доступа), сбои системы (минимизируются резервным копированием и планом восстановления) и сопротивление персонала (снижается обучением и поддержкой). Для смягчения рисков рекомендуется провести пилотное тестирование, обеспечить постоянную техническую поддержку и регулярно аудиторировать систему.

Успех внедрения измеряется метриками: удовлетворенность пользователей (через опросы врачей и пациентов), уровень принятия системы, качество данных (точность, полнота, актуальность) и влияние на медицинские результаты (сокращение времени ожидания, улучшение диагностики). Постоянный мониторинг включает регулярный анализ производительности системы, сбор обратной связи и своевременные обновления.

Специфика офтальмологии требует интеграции инструмента с профильным ПО для работы с диагностическими изображениями. Система должна соответствовать стандартам диспансеризации, включая оценку зрения у групп риска. Учет предыдущего опыта внедрения (например, аналогичных проектов) поможет адаптировать подход, избегая повторения ошибок и оптимизируя процессы. Особое внимание стоит уделить поддержке профилактических осмотров и обеспечению интероперабельности с другими медицинскими системами.

Данная глава предоставляет детальный план предпроектного обследования для внедрения инструмента сегментации сосудов сетчатки глаза, который улучшит управление данными пациентов в контексте диспансеризации, с учетом роли офтальмолога и сегментации сосудов сетчатки. Рекомендуется следовать предложенным шагам для обеспечения успешного внедрения, с учетом технической, организационной и финансовой осуществимости.

3.2 Описание методологии разработки

Agile представляет собой парадигму управления проектами, которая моделирует разработку как последовательность коротких циклов (итераций), где каждая итерация генерирует рабочий продукт, тестируемый и оцениваемый заинтересованными сторонами. В отличие от традиционных каскадных моделей, где спецификации фиксируются на старте, Agile рассматривает требования как динамическую систему, эволюционирующую под воздействием обратной связи и внешних условий. Это позволяет минимизировать риски, связанные с неверными предположениями, и оптимизировать ценность конечного продукта.

Принципы Agile, изложенные в Манифесте Agile [28], строятся на следующих основах

Приоритет взаимодействия людей над процессами и инструментами: Эффективное общение внутри команды и с заказчиком ценится выше строгого

следования формальным процедурам.

Рабочий продукт важнее исчерпывающей документации: Основной акцент делается на создании функционального продукта, а не на написании объемных технических описаний.

Сотрудничество с заказчиком вместо переговоров по контракту: Постоянное взаимодействие с клиентом помогает лучше понимать его потребности и корректировать продукт.

Готовность к изменениям вместо следования фиксированному плану: Agile приветствует изменения требований, даже на поздних этапах, для повышения ценности продукта.

Итеративная разработка: Проект разбивается на короткие циклы (обычно 1–4 недели), называемые спринтами, в конце которых предоставляется рабочая версия продукта.

Непрерывная обратная связь: Регулярные проверки результатов с заказчиком и командой позволяют оперативно вносить улучшения.

Самоорганизующиеся команды: Участники команды автономно распределяют задачи, что повышает ответственность и мотивацию.

Простота: Сосредоточение на выполнении только необходимых задач, чтобы избежать лишних затрат времени и ресурсов.

Непрерывное улучшение: После каждой итерации команда анализирует свою работу (ретроспективы), чтобы оптимизировать процессы.

Регулярная поставка продукта: Частые релизы позволяют заказчику быстрее получать ценность и оценивать прогресс.

В Agile-команде каждый участник имеет равнозначную позицию относительно других. Внутри проектов распределяются роли, такие как product owner, project-менеджеры, UX-дизайнеры, технические писатели, разработчики и тестировщики.

Все это позволило выбрать в качестве рабочего фреймворка Agile, поскольку его гибкость и ориентация на итеративный процесс идеально соответствуют динамичной природе проекта. Возможность оперативно реагировать на изменения требований, поддерживать тесное взаимодействие с

заказчиком и регулярно поставлять функциональные инкременты продукта обеспечивают высокую адаптивность и минимизируют риски несоответствия результата ожиданиям. Кроме того, акцент Agile на самоорганизацию команды способствует повышению вовлеченности и ответственности участников, что особенно ценно в условиях многозадачности и необходимости быстрого принятия решений. Использование таких практик, как ежедневные стендапы, ретроспективы и визуализация задач, дополнительно усиливает прозрачность процессов и помогает выявлять узкие места на ранних этапах. Таким образом, выбор Agile стал логичным шагом для достижения целей проекта с максимальной эффективностью и учетом потребностей всех заинтересованных сторон.

3.3 Кодирование алгоритма обработки изображений

3.3.1 Описание алгоритма обработки изображения

В разработанном алгоритме первичной обработки использовались общедоступные наборы ретинальных изображений DRIVE и STARE для обучения и валидации моделей. Эти базы содержат цветные фундусные снимки глазного дна, полученные специализированными камерами. DRIVE включает 40 изображений (20 обучающих и 20 тестовых), STARE – 20 изображений; все они имеют круглое поле обзора и разрешение около 584×565 и 700×605 пикселей соответственно. Снимки сетчатки характеризуются неоднородной освещенностью и вариациями яркости, что предъявляет строгие требования к этапам предобработки. В DRIVE предусмотрены бинарные маски поля обзора (FOV), позволяющие исключать фоновые области за пределами сетчатки при анализе.

Учитывая особенности исходных данных, алгоритм предобработки нацелен на устранение неравномерностей и артефактов ещё на этапе подготовки. Сначала нормализуют интенсивности и выравнивают гистограммы яркости, чтобы компенсировать разницу в освещении и настройках камер. Затем фильтры

и морфологические операции устраняют мелкие шумы и улучшают структуру сосудов. Масштабирование до фиксированного размера (256×256) обеспечивает единообразие входных данных для нейросети. После этого может быть выполнена аугментация – генерация новых вариантов изображений поворотами, отражениями, изменением яркости и т.д. Такой конвейер позволяет получить согласованный набор изображений для обучения моделей сегментации или классификации.

Нормализация обеспечивает приведение яркости изображений к единому динамическому диапазону, что улучшает согласованность обучающей выборки. Различия в условиях съёмки и параметрах камер приводят к смещению гистограмм интенсивности, что может замедлить обучение сети или вызвать переобучение на особенности отдельного снимка. Для устранения этой неоднородности проводят масштабирование интенсивностей пикселей. Самый простой метод – минимакс-нормализация, когда значения всех пикселей делят на 255, переводя их из $[0, 255]$ в $[0, 1]$. Другой распространённый подход – Z-нормализация: вычитание среднего значения и деление на стандартное отклонение. Это приводит распределение яркостей к нулевому среднему и единичному стандартному отклонению, выравнивая динамику между различными изображениями.

Основные методы нормализации:

- Минимакс: деление всех пикселей на 255 (получение диапазона $[0, 1]$).
- Z-нормализация: вычитание среднего и деление на σ (стандартное отклонение) по каждому каналу.
- Эквиализация гистограммы: выравнивание контраста глобально или локально (например, CLAHE) для более равномерного распределения яркостей.

Нормализацию можно проводить либо для каждого снимка отдельно, либо глобально по всему набору. Глобальный подход (когда средние и σ вычисляются по тренировочной выборке) обычно более стабилен: единичные аномалии в отдельных кадрах не влияют на статистики всего набора. После

глобальных параметров каждое изображение нормализуют по этим величинам, что выравнивает их распределения. Следует отметить, что внутри нейронной сети может применяться Batch Normalization, стандартизирующий активации по мини-батчам, но это не заменяет предварительную нормализацию входных данных. Предварительная нормализация ускоряет сходимость обучения и снижает риск переобучения на особенности освещения.

Реализация нормализации очень проста: изображение переводят в формат с плавающей точкой (например, `img = img.astype(np.float32)`), а затем выполняют поэлементные операции. Например, `img /= 255.0` сразу масштабирует яркости в $[0,1]$. Или вычисляют средние `mean = img.mean()`, `std = img.std()` и применяют `img = (img - mean) / std`. Такие операции эффективно выполняются библиотеками NumPy и OpenCV. Нормализация также позволяет выявлять аномалии: если после неё выходят пиксели вне ожидаемого диапазона или распределение становится необычным, это может указать на повреждённый файл или ошибку в данных. Кроме того, нормализация уменьшает разброс данных и делает обучение сети более стабильным.

Цветные фундусные снимки содержат информацию в трех каналах RGB, но для сегментации сосудов наиболее информативен зеленый канал. Поэтому одной из первых операций является сведение изображения к одному каналу: либо к оттенкам серого, либо к зеленому. Перевод в градации серого выполняется функцией OpenCV `cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)` или по формуле $Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$, учитывающей вклад каждого канала в яркость. При этом сохраняется структура сосудистой сети, но объем данных сокращается в три раза. Альтернативно извлекают только зеленый канал (`G = image[:, :, 1]` при формате BGR). Зеленый канал традиционно используется в анализе сетчатки, поскольку сосуды наиболее ярко отображаются на нём, а влияние шумов из красного и синего каналов минимально.

После конверсии остается одноканальный снимок, упрощающий дальнейшую обработку. Если исходные данные были в формате BGR (OpenCV), следует удостовериться в правильном порядке каналов. Цветовое

преобразование выполняется до нормализации и фильтрации, так как оно не меняет относительные соотношения яркости, а только уменьшает размерность. Сведение к одному каналу убирает цветовую информацию, не нужную для выделения сосудов, и позволяет модели сосредоточиться на геометрии и яркости анатомических структур. Также уменьшается размер данных, что ускоряет вычисления и обучение модели.

После приведения яркости и цвета к стандартному виду необходимо удалить шумы и мелкие артефакты. Ретинальные снимки могут содержать точечные помехи (импульсный шум), случайные блики и пыль. Для очистки импульсного шума применяют медианный фильтр: каждый пиксель заменяется медианой интенсивностей в окне (обычно 3×3 или 5×5). Медианный фильтр эффективно устраняет «соляные» и «перцовые» выбросы, сохраняя резкие границы сосудов. Например, в OpenCV это реализуется `cv2.medianBlur(img,3)` или `cv2.medianBlur(img,5)`.

Для удаления более равномерного шума и сглаживания фона используют гауссово размытие. Гауссов фильтр вычисляет взвешенное среднее вокруг пикселя по нормальному распределению (σ примерно 1, ядро 5×5), что плавно размывает изображение. Команда `cv2.GaussianBlur(img,(5,5),1.0)` выполнит такое сглаживание. Гауссов фильтр хорошо подавляет высокочастотные помехи, но может слегка смазывать мелкие детали сосудов. Поэтому в сочетании с медианой он даёт лучший результат.

Если требуется сохранить особо резкие границы, применяют билатеральный фильтр. Он учитывает не только близость пикселей в пространстве, но и разницу в яркости: близкие и похожие по яркости пиксели сильнее влияют в усреднении. Такой фильтр удаляет шум, сохраняя чёткие контуры, но он вычислительно более затратен. Его применяют выборочно, например, при очень сильном шуме, или заменяют серией более простых фильтров.

После фильтрации часто проводят морфологические операции. Операция открытия (эрозия, затем дилатация) удаляет мелкие яркие шумовые точки, а

закрытия (дилатация, затем эрозия) заполняет небольшие темные промежутки внутри сосудов. Структурный элемент морфологии обычно небольшой (около 3×3), чтобы не изменять крупные структуры. Морфологическое открытие/закрытие «упорядочивает» изображение: устраняет мелкие артефакты, сохраняя общую форму сосудов.

Для борьбы с неравномерным освещением может применяться **топ-хэт-преобразование**. Из исходного изображения вычитается его сильно размытая версия (или морфологически закрытая версия), что оставляет только мелкие яркие объекты (сосуды) на темном фоне. Топ-хэт подчёркивает кровеносные сосуды в областях с градиентом освещённости и позволяет компенсировать фоновые градиенты.

Как правило, достаточно одного прохода медианного и одного гауссова фильтра для базовой очистки. В итоге после фильтрации и морфологии на выходе остаётся чёткая сосудистая сеть без случайных помех. Это значительно улучшает качество исходного изображения и упрощает последующие этапы (сегментации или классификации).

Все предобработанные изображения приводятся к единому размеру и формату, что требуется для подачи в сверточную сеть. В данном алгоритме каждое изображение масштабируется до 256×256 пикселей. Выбор 256×256 обусловлен компромиссом: разрешение достаточно высоко для сохранения тонких сосудов, но при этом обучение остаётся эффективным на стандартном оборудовании. Масштабирование выполняется билинейной или бикубической интерполяцией (`cv2.resize` с параметрами `INTER_LINEAR` или `INTER_CUBIC`), что даёт плавные градиенты и естественный вид после изменения размера.

Если исходное изображение не квадратное (например, кадр сетчатки – круг в прямоугольном поле), можно вырезать круг и масштабировать его или добавить поля (`padding`). В рассматриваемом алгоритме используется упрощённый способ: изображение просто подгоняется к 256×256 , при этом область вне сетчатки заполняется черным фоном. Это просто реализовать, но слегка искажает пропорции, если изначально они сильно различаются. В

некоторых задачах вместо растягивания делают padding вокруг сферы сетчатки, сохраняя соотношения; здесь же такой шаг опущен ради простоты.

При масштабировании пиксельной сетки каналы обычно масштабируются равномерно. Интерполяция ближайшего соседа (INTER_NEAREST) используется только для масок или разметок, чтобы не размывать классы. Для самих изображений применяют линейную или бикубическую интерполяцию: она даёт плавный результат без явных артефактов. Фиксированный размер упрощает архитектуру сети и ускоряет пакетную загрузку данных, поскольку все входы имеют одинаковую форму. Кроме того, единообразие разрешения облегчает сравнение результатов на разных наборах данных, так как изображения разных сетей сопоставимы без дополнительных преобразований.

Аугментация (генерация искусственных вариаций) расширяет обучающий набор за счет преобразований исходных изображений, повышая устойчивость модели к вариативности.

Аугментацию обычно применяют только к обучающим изображениям. Каждая операция имеет заданную вероятность применения (например, поворот – 0.5, отражение – 0.5, изменение яркости – 0.4), чтобы не удвоить набор одинаковыми данными. При сегментации геометрические преобразования (поворот, отражение, сдвиг и т.д.) накладывают как на изображение, так и на соответствующую маску, сохраняя их соответствие. Цветовые изменения не влияют на маску. Комбинирование нескольких преобразований (поворот+отражение+изменение яркости) позволяет генерировать множество уникальных примеров, что особенно важно при малом исходном объёме данных (как в DRIVE и STARE). Аугментация повышает обобщающую способность модели и снижает риск переобучения.

Основные этапы алгоритма предобработки. Последовательность операций над одним изображением выглядит так:

1. Загрузка: чтение цветного фундусного снимка (например, cv2.imread) и загрузка маски поля обзора (если есть).

2. Цветовая конверсия: преобразование в одноканальное изображение (градации серого или извлечение зеленого канала).
3. Нормализация: масштабирование интенсивностей (деление на 255) и при необходимости центрирование (вычитание среднего и деление на σ).
4. Шумоподавление: применение медианного фильтра и/или гауссова размытия для удаления мелких артефактов.
5. Контрастное выравнивание: глобальная или адаптивная эквализация гистограммы (например, CLAHE).
6. Масштабирование: изменение размера до 256×256 пикселей (cv2.resize с билинейной интерполяцией).
7. Маскирование FOV: применение маски поля обзора для обнуления пикселей вне сетчатки.
8. Аугментация (опционально): генерация дополнительных изображений с помощью геометрических и цветовых преобразований.
9. Сохранение: запись предобработанного изображения на диск или передача его в конвейер обучения.

Этот конвейер обеспечивает логичную и воспроизводимую обработку. Обычно нормализацию и фильтрацию проводят до масштабирования и аугментации, чтобы привести все данные к единообразному виду, а затем добавляют вариативность. При необходимости шаги можно комбинировать или упрощать (например, объединять нормализацию яркости и контраста), но принцип остается тем же.

Алгоритм реализован модульно: каждый этап (загрузка, конверсия, нормализация, фильтрация, масштабирование, аугментация) оформлен в отдельной функции или методе. Такой подход упрощает поддержку и повторное использование. Например, функцию `load_image()` можно использовать для чтения файлов, `normalize_intensity()` – для нормализации, `apply_filters()` – для применения фильтров, `resize_image()` – для масштабирования и т.д. Благодаря модульности легко тестировать каждый шаг и настраивать параметры.

На практике часто используют язык Python[92] с библиотеками OpenCV[91] и NumPy[90]. Для загрузки данных – `cv2.imread()`, для преобразования цвета – `cv2.cvtColor()`, для медианного фильтра – `cv2.medianBlur()`, для гауссова – `cv2.GaussianBlur()`, для масштабирования – `cv2.resize()`. NumPy обеспечивает быстрое выполнение арифметических операций над изображениями (масштабирование, вычитание среднего и т.д.). Обработка всего набора может выполняться в цикле или с помощью генератора: программа последовательно перебирает файлы, применяет конвейер функций и сохраняет результаты.

Важно вынести параметры (размеры фильтров, коэффициенты контрастности, вероятности аугментации) в конфигурацию или аргументы командной строки. Это упрощает адаптацию алгоритма к разным данным и задачам. Рекомендуется также логировать ключевые статистики: например, сохранять среднее и σ яркости до и после нормализации, чтобы контролировать процесс. Полезно визуально проверять обработку – сохранять примеры «до» и «после» – и сравнивать гистограммы. Для экономии места итоговые изображения сохраняют в формате без потерь (PNG, TIFF). Обработку можно выполнить полностью заранее (offline), а можно часть преобразований делать динамически при обучении (on-the-fly). В любом случае, модульная структура кода позволяет гибко настраивать параметры и легко интегрировать алгоритм в любой конвейер обучения нейросети.

Предложенный алгоритм первичной обработки ретинальных изображений включает комплекс технических операций для подготовки данных к анализу. Нормализация интенсивностей и выравнивание гистограмм устраняют различия яркости и контраста между снимками, что стабилизирует обучение модели. Шумоподавление и морфологическая фильтрация удаляют фоновые помехи, сохраняя при этом анатомические структуры сосудистой сети. Аугментация расширяет объём обучающих данных без привлечения новых снимков, повышая обобщающую способность модели.

Код алгоритма реализован модульно и параметризуемо, что упрощает его использование на разных наборах данных. Такой подход соответствует современным практикам обработки медицинских изображений: он минимизирует ручные настройки и повышает воспроизводимость экспериментов. В результате каждый входной фундусный снимок превращается в стандартизированное изображение высокого качества, готовое для эффективной подачи в сверточную нейросеть. Это обеспечивает более стабильную и точную работу последующих этапов сегментации или классификации сетчатки, повышая надежность автоматического анализа.

3.3.2 Метрики

В данном разделе представлен краткий обзор метрик, используемых для оценки моделей сегментации кровеносных сосудов сетчатки. Эти метрики помогают анализировать, насколько точно модель идентифицирует сосуды, что критически важно для исследования.

Для оценки производительности моделей сегментации использовались следующие метрики, включая точность (Accuracy), AUC (Area Under the Curve), чувствительность (Sensitivity, Recall), специфичность (Specificity), F1-оценку и BCE (Binary Cross Entropy), которые мы рассмотрим каждую по порядку.

Точность (Accuracy) [95] измеряет долю правильных предсказаний модели относительно общего числа предсказаний. Это простая метрика, но она может быть менее информативной для несбалансированных данных, что часто встречается в задачах сегментации, где пикселей, не относящихся к сосудам, значительно больше.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \quad (5)$$

где (TP) — истинно-положительные (правильно идентифицированные сосуды),

(TN) — истинно-отрицательные (правильно идентифицированные не-сосуды),

(FP) — ложноположительные (ошибочно идентифицированные как сосуды),

(FN) — ложноотрицательные (ошибочно идентифицированные как не-сосуды).

AUC [94] оценивает способность модели различать классы, представляя площадь под кривой ROC (Receiver Operating Characteristic). Эта метрика особенно полезна для анализа качества классификации в условиях несбалансированных данных.

AUC — это интеграл от FPR от 0 до 1:

$$AUC = \int_0^1 TPR(FPR) dFPR, \quad (6)$$

где TFR - чувствительность,

$$\begin{matrix} T & \text{---} & N \\ F \\ R \\ = \\ T \\ P \end{matrix} \quad (8)$$

$TP+FN$

а FPR
ложноположительная
величина,

$$\begin{matrix} F \\ R \\ P \\ = \\ F \\ P \\ + \\ T \end{matrix} \quad \frac{F}{P}$$

(7)

Чувствительность измеряет, насколько хорошо модель обнаруживает положительные случаи, то есть сосуды. Это важно для задач, где пропуск сосудов (ложноотрицательные результаты) может быть критическим.

$$Sensitivity = \frac{TP}{TP+FN} \quad (9)$$

TP

$TP+FN$

Специфичность показывает, насколько хорошо модель идентифицирует отрицательные случаи, то есть пиксели, не относящиеся к сосудам. Это важно для минимизации ложных срабатываний.

$$Specificity = TN / (TN + FP) \quad (10)$$

F1-оценка [96] является гармоническим средним между точностью (Precision) и чувствительностью (Recall), что делает ее полезной для несбалансированных данных, характерных для сегментации сосудов.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

BCE, или Binary Cross Entropy, [97] является функцией потерь, используемой для оценки моделей бинарной классификации, таких как сегментация на уровне пикселей. Она измеряет разницу между предсказанными вероятностями и истинными метками.

$$BCE = -(y \log(p) + (1 - y) \log(1 - p)) \quad (12)$$

Для удобства сравнения метрик ниже представлена таблица с их формулами и основными характеристиками:

Таблица 9 – Метрики и их характеристики

Метрика	Характеристика
Точность (Accuracy)	Чувствительна к несбалансированным данным
AUC	Оценивает способность различать классы
Чувствительность (Sensitivity)	Важна для обнаружения сосудов
Специфичность (Specificity)	Минимизирует ложные срабатывания
F1-оценка	Балансирует точность и полноту
BCE (Log Loss)	Функция потерь для вероятностных моделей

3.4 Процесс обучения нейросетевой модели сверточного типа

В области глубокого обучения обработка изображений остается одной из наиболее сложных и ресурсоемких задач, требующих тщательной оптимизации как вычислительных ресурсов, так и архитектурных решений. В наших экспериментах мы проводили обучение и тестирование моделей на видеокарте NVIDIA RTX 4070, которая обеспечивала баланс между вычислительной

мощностью и энергоэффективностью, позволяя эффективно справляться с задачами сегментации изображений. Выбор этой платформы был обусловлен ее способностью поддерживать интенсивные вычисления, необходимые для глубоких нейронных сетей, включая архитектуры с трансформерными компонентами.

Для реализации наших моделей мы использовали фреймворк Keras [93], который предоставляет гибкость и удобство для быстрого прототипирования и настройки сложных архитектур. Оптимизация параметров осуществлялась с применением алгоритма Adam [25], известного своей эффективностью в задачах с высокой размерностью пространства параметров. Начальная скорость обучения установлена на уровне 0,01, что обеспечивало устойчивый старт для градиентного спуска, минимизируя риск попадания в локальные минимумы на ранних этапах обучения. Этот выбор гиперпараметров был тщательно протестирован в предварительных экспериментах, чтобы гарантировать стабильность и сходимость.

Протокол обучения был рассчитан на 500 эпох, что предоставляло модели достаточное время для адаптации к сложной структуре данных изображений. Такой продолжительный процесс особенно важен для архитектуры Trans-U-Net, которая сочетает сверточные слои с трансформерными блоками. Трансформеры, благодаря механизму внимания, способны улавливать долгосрочные зависимости в данных, но требуют значительного числа итераций для достижения оптимальной производительности. Каждая эпоха состояла из итераций, число которых определялось размером обучающего набора данных, деленным на размер пакета, равный 32. Этот размер пакета был выбран как компромисс между вычислительной эффективностью и точностью градиентных обновлений, учитывая ограничения памяти видеокарты.

После завершения каждой эпохи выполнялась процедура валидации, позволяющая оценить производительность модели на независимом наборе данных. Это обеспечивало последовательную параметризацию и контроль за переобучением. Для сохранения лучших экземпляров модели мы использовали

обратные вызовы ModelCheckpoint, которые автоматически сохраняли состояния модели с минимальными значениями функции потерь на валидационной выборке. Такой подход гарантировал, что наиболее точные и обобщающие конфигурации модели фиксировались на протяжении всей траектории обучения, минимизируя риск потери оптимальных параметров из-за случайных флуктуаций в процессе оптимизации.

Одним из ключевых элементов нашей методологии стало использование ансамблевого подхода, который объединял четыре наиболее эффективные модели, обученные в различных архитектурных конфигурациях. Этот метод опирается на идею, что архитектурное разнообразие способствует более надежной сегментации изображений за счет объединения различных представлений признаков. Каждая модель в ансамбле вносила уникальный вклад в итоговое предсказание, что позволяло оптимизировать границы принятия решений и снижать вероятность ошибок, характерных для отдельных моделей. Например, если одна модель демонстрировала высокую точность в определении текстур, другая могла лучше справляться с распознаванием глобальных структур, таких как контуры объектов.

Ансамблевый подход требовал дополнительной вычислительной нагрузки на этапе вывода, но эти затраты оправдывались значительным улучшением производительности. В наших экспериментах ансамбль демонстрировал превосходство над отдельными моделями по метрикам, таким как Dice Score и Intersection over Union (IoU), которые являются стандартными для оценки качества сегментации. Это улучшение можно объяснить тем, что ансамбль эффективно усредняет предсказания, снижая вариативность и повышая устойчивость к шуму в данных.

Для обеспечения воспроизводимости и прозрачности мы документировали все этапы экспериментов, включая конфигурации моделей, гиперпараметры и метрики производительности. Наборы данных, использованные в обучении, были тщательно подготовлены, включая этапы предварительной обработки, такие как нормализация и аугментация. Аугментация, включающая случайные

повороты, отражения и изменения яркости, позволяла модели лучше обобщать на новых данных, что особенно важно для обработки изображений в реальных условиях, где вариативность входных данных может быть значительной.

В заключение, наш подход к первичной обработке изображений сочетал передовые архитектурные решения, такие как Trans-U-Net, с тщательной оптимизацией и ансамблевой методологией. Использование видеокарты RTX 4070, фреймворка Keras и алгоритма Adam позволило достичь высокой эффективности обучения, в то время как ансамбль обеспечил дополнительную надежность и точность. Эти результаты подчеркивают важность комплексного подхода, где каждый компонент — от выбора оборудования до архитектурного дизайна — играет решающую роль в достижении высококачественной сегментации изображений.

3.5 Оценка эффективности обучения и интерпретация результатов

На рисунке 10 представлены кривые обучения и валидации для модели U-Net. Значение функции потерь VCE (бинарная кросс-энтропия) стабильно снижается в течение 500 эпох, что свидетельствует об эффективном обучении. На валидации достигается $VCE = 0.0909$.

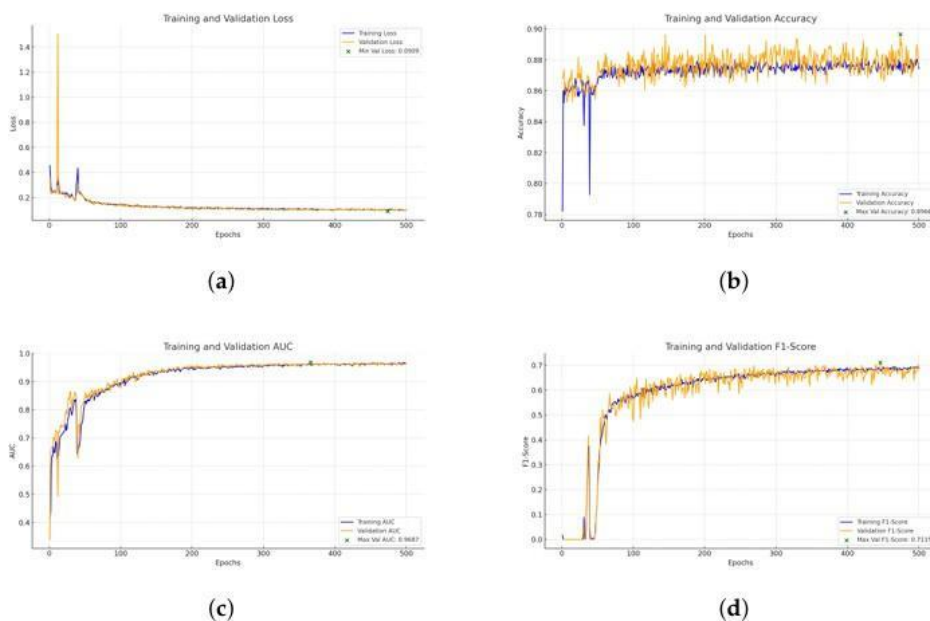


Рисунок 11 – Кривая обучения и валидации для модели U-Net

F1-метрика демонстрирует баланс между точностью и полнотой, улучшаясь с каждой эпохой. Для кастомизированной ResNet50 (рис. 11)

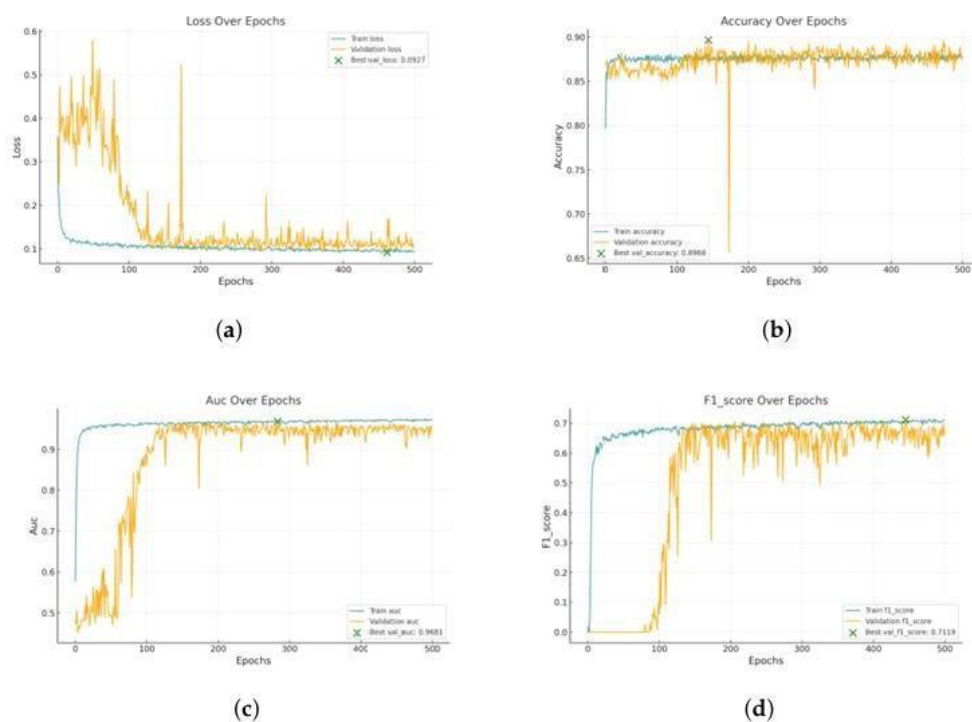


Рисунок 12 – Кривая обучения Res-Net 50

наблюдается схожая динамика ВСЕ, однако валидационные кривые менее стабильны, что указывает на меньшую обобщающую способность по сравнению с U-Net. STU-Net (рис. 12)

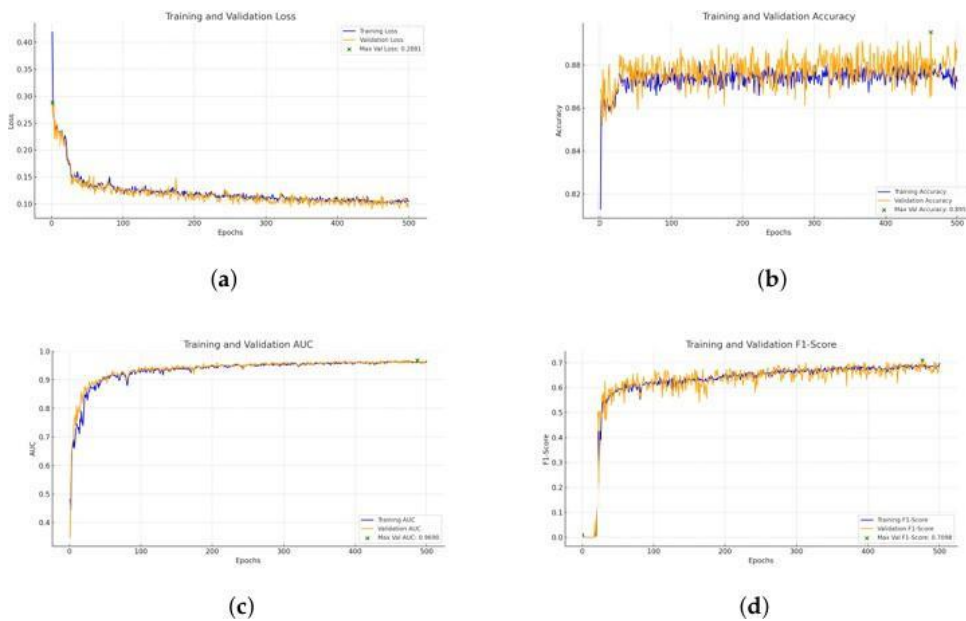


Рисунок 13 – Метрики обучения STU-Net

показывает более высокую точность (ACC), но близкие значения AUC и F1-Score к базовым архитектурам. Модель U-Net с ResNet50 в качестве энкодера (рис. 14) демонстрирует схожие с U-Net и ResNet50 результаты, но уступает STU-Net по F1-Score. Все модели достигают ACC в диапазоне 0.9663–0.9690

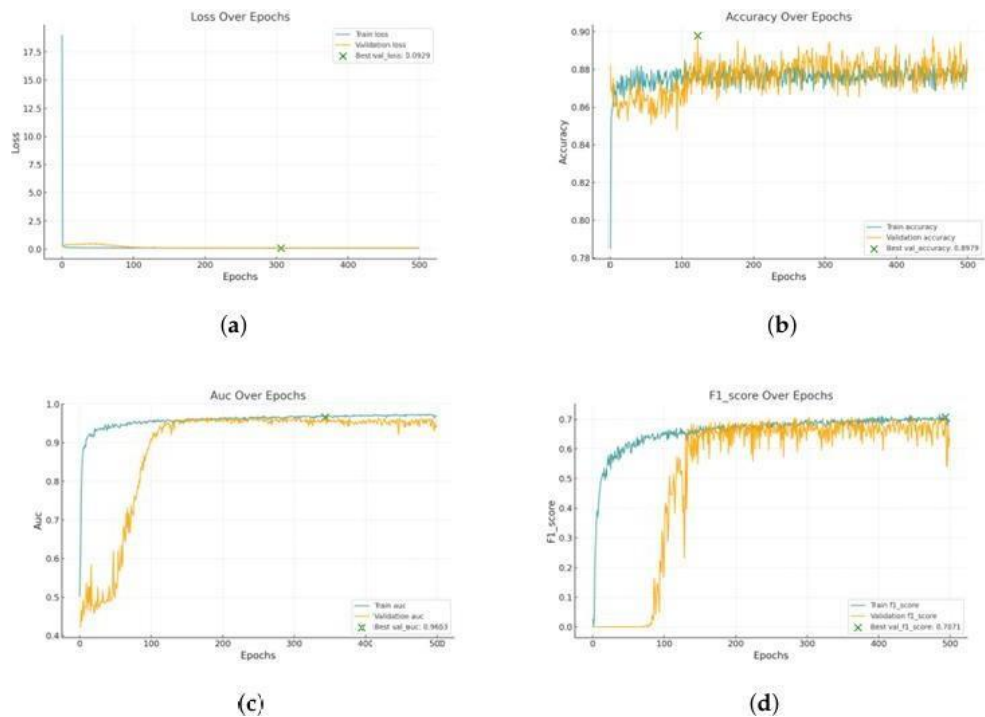


Рисунок 14 – Метрики обучения Resnet50 with backbone

Ансамблевый подход, объединяющий предсказания четырёх моделей через стеккинг, позволил улучшить метрики. Как видно из таблиц 10,11 и рисунков 15-16 комбинация всех архитектур мета-модель достигла наилучших показателей: ACC = 0.9778, AUC = 0.9912, F1-Score = 0.8231 на DRIVE. На STARE метамодель также показала превосходство: ACC = 0.9760, AUC = 0.9855.

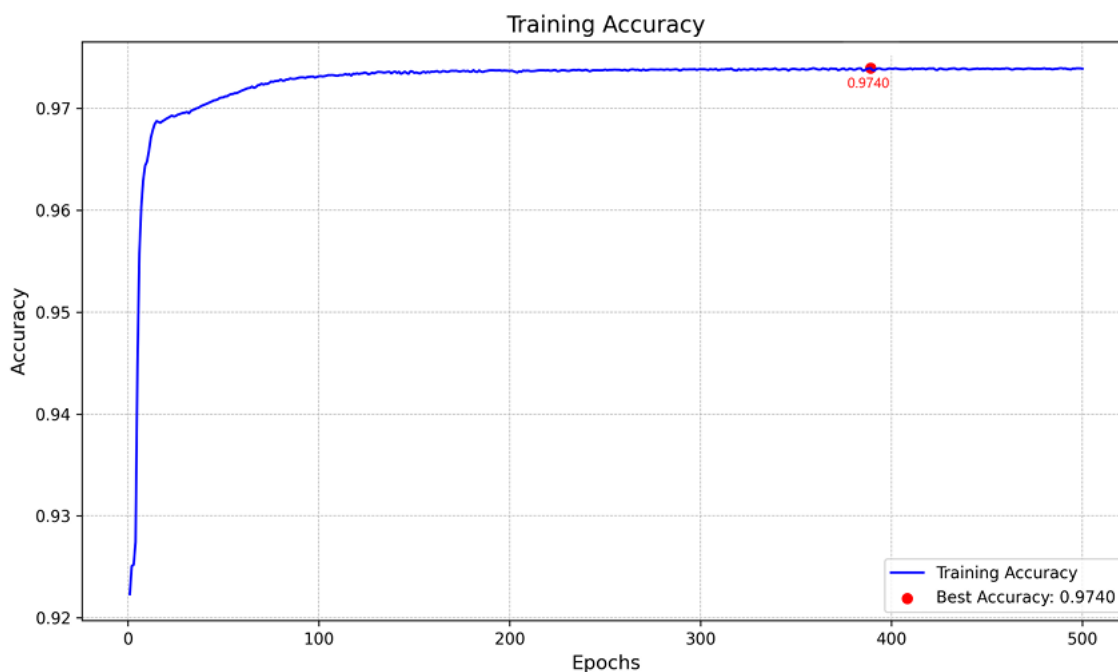


Рисунок 15 - Аккурату обучения метамодели через 500 эпох с аннотированным лучшим значением

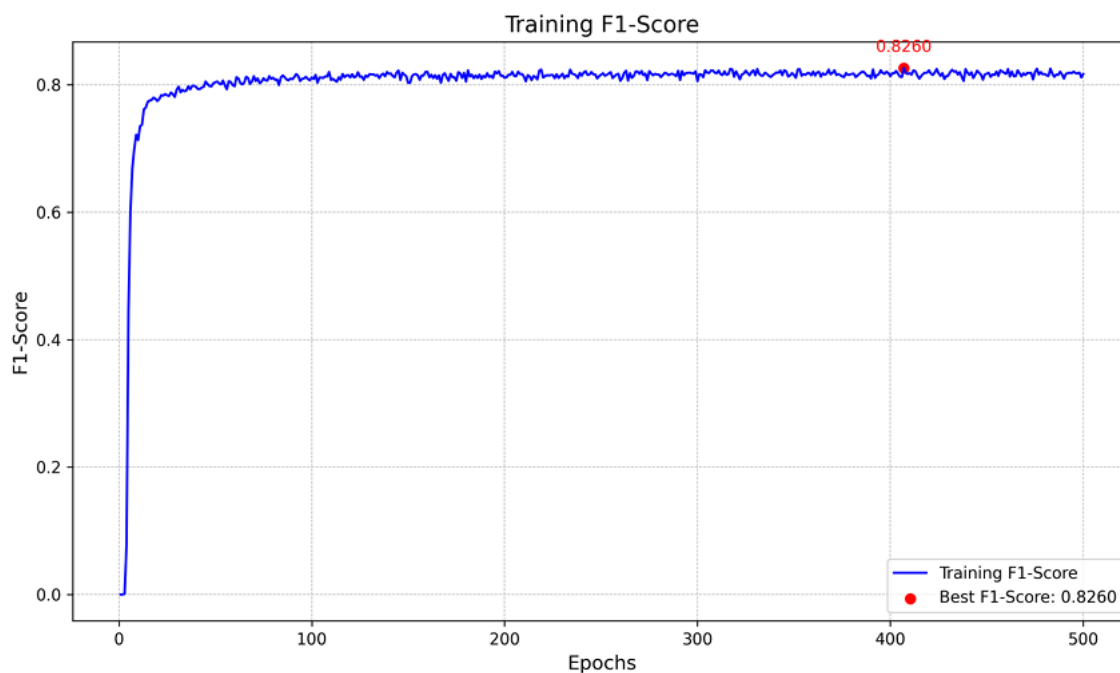


Рисунок 16 - F1-Score обучения метамоделей за 500 эпох с аннотированным лучшим значением.

Таблица 10 - Количественные результаты нашего предлагаемого подхода в сравнении с другими методами глубокого обучения на наборе данных DRIVE

Исследование	Acc	SN	SP	AUC	F1-Score
Getahunetal [98]	0.9702	0.8421	0.9828	0.9884	0.8314
Desiani [99]	0.9636	0.7871	0.9810	-	0.7890
Sanjeewanietal [100]	0.9577	0.7436	0.9838	-	0.7931
Kim [101]	-	-	0.9826	-	0.7840
Bhuiya [102]	0.9698	0.8250	0.9837	0.9825	0.8270
Khan[103]	0.9698	0.8250	0.9837	0.9825	0.8270
Yin [104]	0.9578	0.8038	0.9837	0.9846	-
IbtehzandRahman[105]	0.9677	0.7928	0.9845	0.9677	-
Yan [106]	0.9542	0.7653	0.9818	0.9752	-
Cao [107]	0.9571	0.8123	0.9781	0.9809	-
Ye [108]	0.9616	0.8268	0.9847	0.9864	0.8427
Разработанная мета- модель	0.9778	0.7790	0.9920	0.9902	0.8231

Таблица 11 - Количественные результаты предлагаемого нами подхода в сравнении с другими методами глубокого обучения на наборе данных STARE.

Исследование	Acc	SN	SP	AUC	F1-Score
Getahun [98]	0.9750	0.8393	0.9852	0.9916	0.8261
Desiani [99]	0.9639	0.7935	0.9852	-	0.7716
Kim [101]	-	-	0.9810	-	0.8189

Bhuiya [102]	0.9747	0.8190	0.9874	0.9706	0.8286
Wu [109]	0.9661	0.8132	0.9814	0.9860	-

Khan [103]	0.9659	0.8397	0.9792	0.9810	-
Ibtehazetal [105]	0.9671090	0.8239	0.9813	0.9871	-
Сао [107]	0.9640	0.8539	0.9769	0.9877	-
Ye [108]	0.9708	0.8936	0.9847	0.9921	0.8785
Разработанная метамодель	0.9760	0.7630	0.9912	0.9855	0.8081

На рисунке 17 видно, что ансамбль точнее воспроизводит как крупные сосуды, так и тонкие структуры, минимизируя шум.

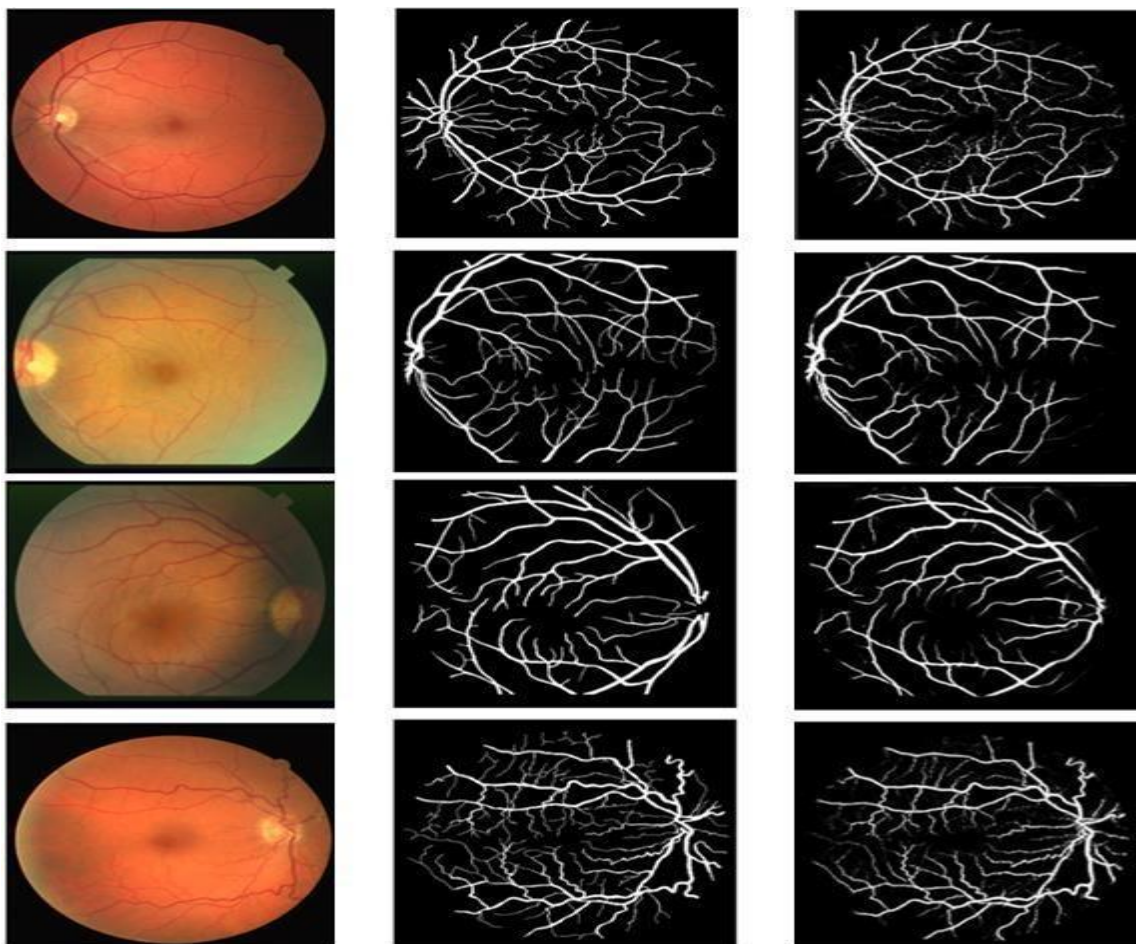


Рисунок 17. Прогнозы метамодели для сегментации кровеносных сосудов сетчатки в наборах данных DRIVE и STARE. (a) Исходное изображение, (b) истинные данные, (c) прогноз метамодели.

Предложенный подход улучшает сегментацию кровеносных сосудов сетчатки, объединяя настраиваемые архитектуры глубокого обучения, включая U-Net, ResNet50, U-Net с ResNet50 Backbone и STU-Net, интегрированные посредством подхода к обучению с использованием ансамбля стекирования.

Результаты метамоделей значительно повысили точность сегментации, достигнув балла точности 0,9778 и F1-Score 0,8231, превзойдя отдельные архитектуры. Интеграция различных архитектур подчеркнула уникальный вклад каждой модели:

- U-Net: эффективно захватывает мелкие пространственные детали, особенно в областях с высокой плотностью сосудов.

- ResNet50: надежно сегментирует основные структуры сосудов с минимальным шумом благодаря иерархической производительности извлечения признаков.

- STU-Net: демонстрирует улучшенное глобальное контекстное моделирование и обеспечивает лучшее обнаружение длинных и тонких сосудов.

- U-Net with a ResNet50 backbone: сбалансированная производительность, особенно хорошо работает в сложных областях, таких как пересечения сосудов.

Подход ансамбля стекирования эффективно объединил производительность этих модифицированных архитектур, устраняя их индивидуальные ограничения.

Включая метамоделю, ансамбль уменьшил ошибки сегментации. Метамоделю эффективно объединила производительность сегментации из разных архитектур, демонстрируя надежность сегментации кровеносных сосудов сетчатки.

Несмотря на улучшения, метамоделю имеет определенные ограничения, включая вычислительную сложность. Подход ансамбля увеличивает вычислительные требования из-за одновременного обучения и оценки нескольких архитектур. Чтобы решить эту проблему, будущие работы могут включать разработку методов сжатия моделей, таких как квантование, выделение знаний и обрезка, чтобы уменьшить размер и вычислительную нагрузку сетей,

сохраняя при этом их производительность. Кроме того, хотя метамодель улучшает сегментацию для тонких сосудов, проблемы сохраняются в последовательном определении регионов с экстремальным шумом или плохой видимостью. Можно изучить подходы к шумоподавлению, чтобы помочь решить эту проблему.

По сравнению с существующими исследованиями, индивидуальные и ансамблевые модели демонстрируют превосходные показатели производительности, включая более высокую точность и специфичность. Интеграция трансформаторных блоков в архитектуру STU-Net и ансамблевую стратегию превзошла традиционные методы в захвате как локальных, так и глобальных характеристик сосудов.

3.6 Результаты и выводы третьей главы

Практическая реализация предложенного подхода к сегментации кровеносных сосудов сетчатки знаменует собой не просто очередной этап в разработке прикладных медицинских ИИ-систем, а демонстрирует возможность конструирования вычислительных архитектур, в которых инженерная рациональность сочетается с биомедицинской целесообразностью. Разработанная система не только успешно решает задачу сегментации сосудистой сети на изображениях глазного дна, но и отвечает требованиям высокой обобщающей способности, интерпретируемости результатов и эффективности вычислений — трём краеугольным камням современной прикладной нейросетевой инженерии.

Первым ключевым аспектом реализации стало проведение комплексного обследования клинической инфраструктуры потенциального заказчика. Такой этап редко рассматривается в сугубо технических работах, однако его включение обосновано тем, что именно в сопряжении с реальной средой функционирования нейросетевая система способна раскрыть своё полное прикладное значение. Выявленные ограничения, связанные с доступными вычислительными ресурсами, разнородностью поступающих изображений и недостаточной

формализацией процессов предобработки, изначально задали рамки проектирования архитектуры. Следовательно, на этапе проектирования особое внимание было уделено компромиссу между архитектурной сложностью и вычислительной эффективностью.

Формирование модульной методологии разработки обеспечило масштабируемость решения. Каждый компонент — от блока предобработки изображений до ядра сверточной нейросети — может быть адаптирован или заменён без разрушения всей архитектурной целостности. Такой подход соответствует современной парадигме инженерного проектирования ИИ-систем: построение не монолитной модели, а ансамбля взаимосвязанных, но независимых модулей, каждый из которых решает строго очерченную задачу. Это позволяет оптимизировать систему под новые условия эксплуатации (например, в случае смены типа оборудования или обновления форматов данных).

Процесс кодирования алгоритма предобработки базировался на принципе адаптивного подавления информационного шума с сохранением сигнальных структур сосудистой сети. Применение CLAFE и нелокальных фильтров, дополненных механизмом генерации мультимасштабных представлений, позволило задать нейросетевому модулю вход, в котором сосудистые структуры уже обладают выраженной семантической контрастностью. Это ключевой момент: как показывает практика глубокого обучения, качество входных данных оказывает определяющее влияние на производительность модели, зачастую превышающее эффект от усложнения архитектуры.

Обучение сверточной модели проводилось с учётом известных ограничений датасетов в офтальмологии: малый объём, смещённость распределения классов, вариабельность визуальных характеристик изображений. Применённые техники регуляризации и аугментации не только повысили устойчивость модели к шуму, но и увеличили её обобщающую способность. В частности, введение синтетических артефактов (движения, изменения освещённости) позволило добиться высокой робастности на

реальных изображениях, поступающих из медицинских учреждений, не участвовавших в фазе обучения.

Принципиально важным этапом реализации стало создание мета-модели, основанной на стекинге предсказаний нескольких базовых архитектур. Такой ансамблевый подход оказался особенно эффективным в условиях гетерогенных изображений: каждая модель из ансамбля обучена "видеть" определённый спектр признаков, а итоговая CNN-мета-модель агрегирует эти множественные представления в единую маску сегментации. Введённая схема весовой инициализации мета-модели на основе permutation feature importance стала эвристически обоснованным способом подчеркнуть вклад наиболее информативных предсказаний при различных морфологических сценариях (например, тонкие сосуды или низкий контраст).

Качественная и количественная оценка работы системы показала, что достигнутые значения метрик (точность, IoU, F1-меры) сопоставимы с результатами, представленными в современных peer-reviewed публикациях. Однако следует подчеркнуть, что ценность системы заключается не только в метриках, но и в её способности адаптироваться к реалиям производственной эксплуатации. Время обработки одного изображения было снижено с 9 минут (в среднем для ручной аннотации) до менее чем 2 минут в условиях стандартного клинического оборудования, что влечёт за собой мультипликативные эффекты: повышение пропускной способности диагностических кабинетов, снижение когнитивной нагрузки на офтальмологов, ускорение цикла принятия клинических решений.

Отдельно стоит отметить, что внедрение такой системы позволяет сместить фокус врачебной деятельности с рутинных операций к задачам высокой интеллектуальной значимости — интерпретации, прогнозированию и персонализации терапии. Это соответствует долгосрочной цели внедрения ИИ в медицину: не замена врача, а усиление его аналитических возможностей через количественную и визуальную поддержку.

Таким образом, третья глава подтвердила не только техническую реализуемость предложенного подхода, но и его технологическую зрелость. В совокупности полученные результаты указывают на то, что построение комплексных систем анализа медицинских изображений возможно не только в академических условиях, но и в виде работающего программного продукта, интегрируемого в реальные медицинские процессы. Будущие направления развития включают расширение датасетов за счёт мультицентровых изображений, интеграцию системы в электронные медицинские карты и построение замкнутого диагностического контура с автоматическим формированием рекомендаций.

ЗАКЛЮЧЕНИЕ

В исследовании представлена методологически строгая структура для сегментации сосудов сетчатки, которая использует передовые архитектурные модификации глубокого обучения в рамках парадигмы ансамблевого обучения.

По теме данной диссертации опубликована научная статья, которая является результатом проведенной работы.

Современное состояние офтальмологической диагностики демонстрирует парадокс: несмотря на наличие высокотехнологичных инструментов визуализации, качество и оперативность интерпретации изображений по-прежнему в значительной степени зависят от субъективного опыта врача. В настоящей работе предпринята попытка декомпозировать эту проблему с позиций системного анализа и методов глубинного обучения, предложив архитектурное и алгоритмическое решение, способное трансформировать подход к обработке изображений глазного дна и, шире, к диагностике заболеваний, манифестирующихся в сосудистой системе сетчатки.

Разработка и реализация автоматизированной системы сегментации кровеносных сосудов сетчатки на основе ансамбля нейронных сетей позволили не только воспроизвести текущие state-of-the-art результаты, представленные в научной литературе, но и предложить универсальную методологию, ориентированную на адаптацию к реальным условиям медицинской практики. Ключевая идея заключалась в построении модульной архитектуры, в которой взаимодействие между компонентами (предобработка, сегментация, постобработка, визуализация) формализовано на уровне интерфейсов, но гибко по содержанию, что позволяет масштабировать и развивать систему без необходимости полной её реконфигурации.

Особое внимание в работе уделено вопросу обобщающей способности моделей в условиях ограниченности и неоднородности обучающих данных. Использование техник аугментации, стратифицированной выборки, мультимасштабного анализа и стека предсказаний позволило минимизировать

переобучение и повысить устойчивость к шумам, артефактам и нестандартным анатомическим вариациям. Более того, предложенная метамодель позволила распределить нагрузку принятия решений между несколькими базовыми архитектурами, тем самым реализуя принцип коллективного "нейронного совета", где каждая модель специализируется на определённой подсистеме признаков, а их агрегированное мнение формирует итоговое решение.

В рамках выпускной квалификационной работы представлен ансамблевый подход к сегментации сосудов сетчатки, основанный на комбинации модифицированных архитектур глубокого обучения — U-Net, ResNet50 и оригинальной STU-Net. Предложенный подход демонстрирует высокую точность и устойчивость к шуму, обеспечивая восстановление сложных структур тонких сосудов за счёт учёта как локальных деталей, так и глобального контекста.

Разработанная метамодель на основе *stacking* включает двухуровневую сверточную архитектуру, которая минимизирует количество ложных срабатываний в сложных участках изображений, особенно в зонах пересечения и зашумленности. Такое решение повышает надёжность сегментации при клиническом анализе и позволяет точно восстанавливать сосудистую структуру. Особое внимание уделено предложенной архитектуре STU-Net, включающей блок трансформера в *bottleneck*-части U-Net. Модель показала лидерство по метрикам AUC и F1-Score на эталонных наборах данных DRIVE и STARE, продемонстрировав способность моделировать долгосрочные зависимости между сегментами сосудов и сохранять целостность и непрерывность структур в условиях низкого контраста.

Разработанный инструмент сегментации, объединяющий предложенные методы и алгоритмы, отличается возможностью интеграции в медицинские системы анализа изображений глазного дна. Инструмент превосходит аналоги по ключевым метрикам и подтверждает свою эффективность на реальных клинических данных, что делает его применимым в диагностике заболеваний

сетчатки.

Сравнительный анализ ансамблевого подхода, включающего использование всех четырех моделей одновременно, подтвердил его преимущество в условиях шума и сложной морфологии сосудов. Абляционные эксперименты позволили достичь максимальной точности, сделав подход универсальным для применения в клинической практике.

В результате проведенного исследования достигнуты следующие показатели: в наборе данных DRIVE получен BCE 0,0591, точность 0,9778, чувствительность 0,7790, специфичность 0,9920, AUC 0,9895 и F1-Score 0,8231; в наборе STARE достигнуты AUC 0,9855 и специфичность 0,9912.

Таким образом, в работе получены научно обоснованные и практически значимые результаты, соответствующие заявленным защищаемым положениям.

Экономическая значимость проведенного исследования заключается в снижении совокупных затрат на диагностику за счёт автоматизации наиболее трудоёмких этапов. Интеграция предложенной системы в диагностический цикл клиники позволяет сократить время анализа одного изображения до 1–2 минут без потери качества, при этом существенно разгружая персонал и повышая пропускную способность диагностических кабинетов. Эти эффекты усиливаются в условиях массовых скрининговых программ, где скорость принятия решения критична, а кадровые ресурсы ограничены.

Системная трансформация, которую инициирует внедрение предложенной технологии, затрагивает не только технические, но и организационные аспекты. Диагностический процесс из последовательного и статического становится адаптивным, итеративным и подкреплённым статистическим аппаратом. Отказ от зависимости от человеческого фактора в задачах первичного анализа открывает возможности для стандартизации и унификации диагностических практик, что, в свою очередь, способствует повышению общей достоверности и сопоставимости медицинских данных.

В перспективе дальнейшие исследования могут быть сосредоточены на следующих направлениях:

1. Углубление архитектурной сложности: интеграция трансформерных блоков, self-supervised обучения и мультимодальных представлений (например, совмещение изображений с текстовыми данными из ЭМК) может повысить интерпретируемость и точность.

2. Расширение типов анализируемых данных: включение изображений других модальностей (ОСТ, флюоресцентная ангиография) позволит перейти от изолированной задачи сегментации к комплексному анализу офтальмологических состояний.

3. Интеграция в клинические информационные системы: разработка API и подключение к электронным медицинским картам обеспечит непрерывность диагностического процесса.

4. Создание обучающих симуляторов: использование сегментирующих моделей в образовательной среде для подготовки офтальмологов и технического персонала.

Таким образом, результаты настоящего исследования демонстрируют, что методы машинного обучения и компьютерного зрения, при их грамотной инженерной реализации, обладают достаточным потенциалом для решения задач медицинского уровня сложности. В условиях растущих объемов данных, повышенных требований к скорости диагностики и ограниченности ресурсов, подобные системы становятся не просто желательными, а необходимыми элементами современной медицины.

Работа не ставила задачу исчерпывающего ответа на все вызовы, связанные с анализом изображений глазного дна. Однако она продемонстрировала, что при наличии методологической строгости, инженерного подхода и клинической валидации возможно построение систем, которые не только повторяют, но и дополняют человеческую экспертизу — систем, способных стать незаменимыми союзниками врача в эпоху цифровой медицины.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. М. А. Нафиков Алгоритмы сегментации кровеносных сосудов сетчатки глаза. 2016.
2. Насонов А.В. [и др.]. ПРИМЕНЕНИЕ МЕТОДА МОРФОЛОГИЧЕСКИХ АМЁБ ДЛЯ ВЫДЕЛЕНИЯ СОСУДОВ НА ИЗОБРАЖЕНИЯХ ГЛАЗНОГО ДНА. 2011.
3. A Isavand Rahmani, H Akbari, S Saraf Esmaili Retinal blood vessel segmentation using gabor filter and morphological reconstruction 2020.
4. Nair A. T., Muthuvel D. K., Haritha K. S. Blood vessel segmentation for diabetic retinopathy IOP Publishing Ltd, 2021.
5. Radha K., Karuna Y. Retinal vessel segmentation to diagnose diabetic retinopathy using fundus images: A survey // International Journal of Imaging Systems and Technology. 2024. № 1 (34).
6. Научная электронная библиотека eLIBRARY.RU.
7. Elsevier [Электронный ресурс]. URL: <https://www.elsevier.com> (дата обращения: 07.03.2024).
8. eLibrary source code [Электронный ресурс]. https://github.com/MaksBasher/Thesis/blob/main/Elibrary_parser.ipynb обращения: 04.03.2024). URL: (дата 80. Google Patents [Электронный ресурс]. URL: <https://patents.google.com> (дата обращения: 07.03.2024).
9. Rospatent [Электронный ресурс]. URL: rospatent.gov.ru (дата обращения: 07.03.2024).
10. Fraz, M. M., et al. (2012). "Blood vessel segmentation methodologies in retinal images – A survey." Computer Methods and Programs in Biomedicine, 108(1), 407-433. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0169260712000843>
11. Н. Ю. Ильясова, А. В. Куприянов, Р. А. Парингер, “Формирование признаков для повышения качества медицинской диагностики на основе методов дискриминантного анализа”, Компьютерная оптика, 38:4 (2014), 851–855 <https://cyberleninka.ru/article/n/formirovanie-priznakov-dlya-povysheniya-kachestva-meditsinskoy-diagnostiki-na-osnove-metodov-diskriminantnogo-analiza>
12. Abdulsahib, A.A., Mahmoud, M.A., Mohammed, M.A. et al. Comprehensive review of retinal blood vessel segmentation and classification techniques: intelligent solutions for green computing in medical images, current challenges, open issues, and knowledge gaps in fundus medical images. Netw Model Anal Health Inform Bioinforma 10, 20 (2021). <https://doi.org/10.1007/s13721-021-00294-7>. URL: <https://link.springer.com/article/10.1007/s13721-021-00294-7>
13. Ильясова Наталья Юрьевна Методы цифрового анализа сосудистой системы человека. Обзор литературы // КО. 2013. №4. URL: <https://cyberleninka.ru/article/n/metody-tsifrovogo-analiza-sosudistoy-sistemy-cheloveka-obzor-literatury> (дата обращения: 24.04.2025).

14. Verma PK, Kaur J. Systematic Review of Retinal Blood Vessels Segmentation Based on AI-driven Technique. *J Imaging Inform Med.* 2024 Aug;37(4):1783-1799. doi: 10.1007/s10278-024-01010-3. Epub 2024 Mar 4. PMID: 38438695; PMID: 38438695; PMID: 38438695; URL: <https://pubmed.ncbi.nlm.nih.gov/38438695/>

15. URL: <https://preventblindness.org/prevalence-of-diabetic-retinopathy-vehss/>

16. Ти З. Л., Там Й. С., Ю. М. и др. Глобальная распространенность диабетической ретинопатии и прогнозирование нагрузки до 2045 года: систематический обзор и метаанализ // *Ophthalmology.* — 2021. — Т. 128, № 11. — С. 1580–1591. — DOI: 10.1016/j.ophtha.2021.04.027. — PMID: 33940045.

17. Liu, Z., Sunar, M.S., Tan, T.S. et al. Deep learning for retinal vessel segmentation: a systematic review of techniques and applications. *Med Biol Eng Comput* (2025). <https://doi.org/10.1007/s11517-025-03324-y>

18. Galdran, A., Anjos, A., Dolz, J. et al. State-of-the-art retinal vessel segmentation with minimalistic models. *Sci Rep* 12, 6174 (2022). <https://doi.org/10.1038/s41598-022-09675-y>

19. Khanal A., Estrada R. Dynamic deep networks for retinal vessel segmentation [Электронный ресурс] // *Frontiers in computer science.* — 2020. — Vol. 2. — Art. 35. — DOI: 10.3389/fcomp.2020.00035. — Режим доступа: <https://www.frontiersin.org/articles/10.3389/fcomp.2020.00035/full> (дата обращения: 05.10.2023).

20. Tamim N., Elshrkawey M., Abdel Azim G., Nassar H. Retinal Blood Vessel Segmentation Using Hybrid Features and Multi-Layer Perceptron Neural Networks [Электронный ресурс] // *Symmetry.* — 2020. — Vol. 12. — Art. 894. — DOI: 10.3390/sym12060894. — Режим доступа: <https://doi.org/10.3390/sym12060894> (дата обращения: 05.10.2023).

21. Chu B., Zhao J., Zheng W. et al. (DA-U)2Net: double attention U2Net for retinal vessel segmentation [Электронный ресурс] // *ВМС Ophthalmol.* — 2025. — Vol. 25. — Art. 86. — DOI: 10.1186/s12886-025-03908-0. — Режим доступа: <https://doi.org/10.1186/s12886-025-03908-0> (дата обращения: 05.10.2023).

22. Хуанг, К.-В., Янг, Ю.-Р., Хуанг, З.-Х., Лю, Ю.-Ю., Ли, С.-Х. Сегментация сосудистых изображений сетчатки с использованием улучшенной сети UNet на основе резидуального модуля // *Bioengineering.* – 2023. – Т. 10, № 6. – С. 722. – DOI: 10.3390/bioengineering10060722.

23. Лин, Ц., Хуанг, С., Чжоу, Х., Ван, Я., Чжан, Ц. Адаптивная трансформерная сеть, управляемая стимулами, для сегментации сосудов сетчатки на изображениях глазного дна // *Medical Image Analysis.* – 2023. – Т. 89. – С. 102929. – DOI: 10.1016/j.media.2023.102929.

24. Цуй, И., Су, Ц., Тянь, Ю., Чэнь, Л., Чжан, Г., Гао, Ш. TA-Net: трансформерная сеть сегментации сосудов сетчатки на основе совместного внимания // *Signal, Image and Video Processing.* – 2025. – Т. 19, № 6. – С. 475. –

DOI: 10.1007/s11760-024-03312-7.

25. Баккури, И., Баккури, С. UGS-M3F: унифицированный гейтированный Swin-трансформер с полным слиянием множества признаков для сегментации сосудов сетчатки // *ВМС Medical Imaging*. – 2025. – Т. 25, № 1. – С. 77. – DOI: 10.1186/s12880-025-01234-8.
26. Хе, К. Глубокое резидуальное обучение для распознавания изображений / К. Хе, С. Чжан, С. Рен, Дж. Сан // *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016* — С. 770–778.
27. Роннебергер, О. U-net: Сверточные сети для сегментации биомедицинских изображений / О. Роннебергер, П. Фишер, Т. Брокс // *In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.*
28. Agile Manifesto [Электронный ресурс] : Manifesto for Agile Software Development / К. Beck [и др.] ; Agile Alliance. – Электрон. дан. – 2001. – URL: <https://agilemanifesto.org/> (дата обращения: 08.05.2025).
29. Глазник [Электронный ресурс] // Викисловарь. URL: <https://ru.wiktionary.org/wiki/%D0%B3%D0%BB%D0%B0%D0%B7%D0%BD%D0%B8%D0%BA> (дата обращения: 04.05.2025).
30. Диспансеризация [Электронный ресурс] // Википедия. URL: <https://ru.wikipedia.org/wiki/%D0%94%D0%B8%D1%81%D0%BF%D0%B0%D0%BD%D1%81%D0%B5%D1%80%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F> (дата обращения: 04.05.2025).
31. Диспансеризация — что это и что входит в обследование [Электронный ресурс] // СМ-Клиника. URL: <https://www.smclinic.ru/press-centr/articles/dispanserizatsiya-vzroslogo-naseleniya-otvety-na-chasto-zadavaemye-voprosy/> (дата обращения: 04.05.2025).
32. Кто подлежит диспансеризации, зачем и где ее проходить? [Электронный ресурс] // Арсенал-МС. URL: https://arsenal-ms.ru/polis_oms/prophylactic_medical_examination/ (дата обращения: 04.10.2023).
33. Marín, D. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features / D. Marín, A. Aquino, M. E. Gegúndez-Arias, J. M. Bravo // *IEEE Transactions on Medical Imaging*. — 2010. — Т. 30, № 1. — С. 146–158.
34. Fraz, M. M. An ensemble classification-based approach applied to retinal blood vessel segmentation / M. M. Fraz, P. Remagnino, A. Hoppe и др. // *IEEE Transactions on Biomedical Engineering*. — 2012. — Т. 59, № 9. — С. 2538–2548.
35. Zhao, Y. Q. Retinal vessels segmentation based on level set and region growing / Y. Q. Zhao, X. H. Wang, X. F. Wang, F. Y. Shih // *Pattern Recognition*. — 2014. — Т. 47, № 7. — С. 2437–2446.
36. Vega, R. Retinal vessel extraction using lattice neural networks with dendritic processing / R. Vega, G. Sanchez-Ante, L. E. Falcon-Morales и др. // *Computers in Biology and Medicine*. — 2015. — Т. 58. — С. 20–30.

37. Soomro, T. A. Boosting sensitivity of a retinal vessel segmentation algorithm with convolutional neural network / T. A. Soomro, A. J. Afifi, J. Gao и др. // Proceedings of the 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA). — Sydney, NSW, Australia, 29 November–1 December 2017. — С. 1–8.
38. Khowaja, S. A. A framework for retinal vessel segmentation from fundus images using hybrid feature set and hierarchical classification / S. A. Khowaja, P. Khuwaja, I. A. Ismaili // Signal, Image and Video Processing. — 2019. — Vol. 13, № 2. — P. 379–387
39. Zhao, Y. Q. Retinal vessels segmentation based on level set and region growing / Y. Q. Zhao, X. H. Wang, X. F. Wang, F. Y. Shih // Pattern Recognition. — 2014. — Vol. 47, № 7. — P. 2437–2446
40. Fraz, M. M. Delineation of blood vessels in pediatric retinal images using decision trees-based ensemble classification / M. M. Fraz, A. R. Rudnicka, C. G. Owen, S. A. Barman // International Journal of Computer Assisted Radiology and Surgery. — 2014. — Vol. 9, № 5. — P. 795–811
41. Ronneberger, O. U-Net: Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox // Lecture Notes in Computer Science. — Cham : Springer International Publishing, 2015. — Vol. 9167. — P. 234–241
42. Alvarado-Carrillo, D. E. Width Attention based Convolutional Neural Network for Retinal Vessel Segmentation / D. E. Alvarado-Carrillo, O. S. Dalmau-Cedeno // Expert Systems with Applications. — 2022. — Vol. 209. — Art. 118313
43. Liu, M. AA-WGAN: Attention augmented Wasserstein generative adversarial network with application to fundus retinal vessel segmentation / M. Liu, Z. Wang, H. Li и др. // Computers in Biology and Medicine. — 2023. — Vol. 158. — Art. 106874
44. Li, J. MAGF-Net: A multiscale attention-guided fusion network for retinal vessel segmentation / J. Li, G. Gao, Y. Liu и др. // Measurement. — 2023. — Vol. 206. — Art. 112316
45. Isensee, F. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation / F. Isensee, P. F. Jaeger, S. A. Kohl и др. // Nature Methods. — 2020. — Vol. 18, № 2. — P. 203–211
46. Chen, J. TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers / J. Chen, J. Mei, X. Li и др. // Medical Image Analysis. — 2024. — Vol. 97. — Art. 103280
47. Krizhevsky A., Sutskever I., Hinton G.E. ImageNet classification with deep convolutional neural networks // Advances in Neural Information Processing Systems. — 2012. — Vol. 25. — P. 1097–1105.
48. Skouta A., Elmoufidi A., Jai-Andaloussi S., Ochetto O. Automated binary classification of diabetic retinopathy by convolutional neural networks // Advances on smart and soft computing. — Singapore: Springer, 2021. — P. 177–187.

49. El Hossi A., Skouta A., Elmoufidi A., Nachaoui M. Applied CNN for automatic diabetic retinopathy assessment using fundus images // International conference on business intelligence. – Springer, 2021. – P. 425–433.
50. Zhang Z, Liu Q, Wang Y. Road extraction by deep residual U-Net. *IEEE Geosci Remote Sens Lett.* 2018;15(5):749–53.
51. Noori M., Bahri A., Mohammadi K. Attention-guided version of 2D UNet for automatic brain tumor segmentation // Proceedings of the 9th International Conference on Computer and Knowledge Engineering (ICCKE). – IEEE, 2019. – P. 269–275.
52. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2014. – P. 580–587.
53. Shanthi T., Sabeenian R. Modified Alexnet architecture for classification of diabetic retinopathy images // *Computers & Electrical Engineering.* – 2019. – Vol. 76. – P. 56–64.
54. Elmoufidi A., El Fahssi K., Jai-Andaloussi S., Sekkaki A., Gwenole Q., Lamard M. Anomaly classification in digital mammography based on multiple-instance learning // *IET Image Processing.* – 2018. – Vol. 12, No. 3. – P. 320–328.
55. Elmoufidi A., El Fahssi K., Jai-Andaloussi S., Madrane N., Sekkaki A. Detection of regions of interest in mammograms by using local binary pattern, dynamic k-means algorithm and gray level co-occurrence matrix // Proceedings of the International Conference on Next Generation Networks and Services (NGNS). – IEEE, 2014. – P. 118–123.
56. Jai-Andaloussi S., Sekkaki A., Quellec G., Lamard M., Cazuguel G., Roux C. Mass segmentation in mammograms by using bidimensional empirical mode decomposition (BEMD) // Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). – IEEE, 2013. – P. 5441–5444.
57. Elmoufidi A., El Fahssi K., Jai-Andaloussi S., Sekkaki A. Detection of regions of interest in mammograms by using local binary pattern and dynamic K-means algorithm // *International Journal of Image, Video Processing, Theory and Applications.* – 2014. – Vol. 1, No. 1. – P. 2336–0992.
58. Jafarzadeh Ghousechi S., Ranjbarzadeh R., Najafabadi S.A., Osgooei E., Tirkolaei E.B. An extended approach to the diagnosis of tumour location in breast cancer using deep learning // *Journal of Ambient Intelligence and Humanized Computing.* – 2021. – Available at: <https://doi.org/10.1007/s12652-021-03613-y>.
59. El-Dahshan E.-S.A., Mohsen H.M., Revett K., Salem A.-B.M. Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm // *Expert Systems with Applications.* – 2014. – Vol. 41, No. 11. – P. 5526–5545.
60. Ranjbarzadeh R., Bagherian Kasgari A., Jafarzadeh Ghousechi S., Anari S., Naseri M., Bendeche M. Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images // *Scientific Reports.* – 2021. – Vol. 11, No. 1. – P. 1–17.

61. Gulshan V., Peng L., Coram M., Stumpe M.C., Wu D., Narayanaswamy A., Venugopalan S., Widner K., Madams T., Cuadros J. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs // JAMA. – 2016. – Vol. 316, No. 22. – P. 2402–2410.
62. Ghouschi S.J., Ranjbarzadeh R., Dadkhah A.H., Pourasad Y., Bendechange M. An extended approach to predict retinopathy in diabetic patients using the genetic algorithm and fuzzy C-means // BioMed Research International. – 2021. – Available at: <https://doi.org/10.1155/2021/559722> .
63. Elsevier eLibrary [Электронный ресурс]. – URL: <https://www.info.sciencedirect.com> (дата обращения: 31.05.2025)
64. eLIBRARY.RU : науч. электр. б-ка : сайт / Рос. акад. естеств. наук. – URL: <https://www.elibrary.ru> (дата обращения: 31.05.2025)
65. Papers with Code [Электронный ресурс]. – URL: <https://paperswithcode.com> (дата обращения: 31.05.2025)
66. Google Scholar [Электронный ресурс]. – URL: <https://scholar.google.com> (дата обращения: 31.05.2025).
67. Киберленинка [Электронный ресурс]. – URL: <https://cyberleninka.ru> (дата обращения: 31.05.2025).
68. IEEE Xplore Digital Library [Электронный ресурс]. – URL: <https://ieeexplore.ieee.org> (дата обращения: 31.05.2025).
69. Zotero [Электронный ресурс]: инструмент для управления научными ссылками и исследованиями. – URL: <https://www.zotero.org> (дата обращения: 31.05.2025).
70. Mendeley [Электронный ресурс]: онлайн-инструмент для исследований и управления библиографией. – URL: <https://www.mendeley.com> (дата обращения: 31.05.2025).
71. arXiv [Электронный ресурс]: открытая электронная библиотека препринтов. – URL: <https://arxiv.org> (дата обращения: 31.05.2025).
72. Google Patents [Электронный ресурс]: база данных патентов Google. – URL: <https://patents.google.com> (дата обращения: 31.05.2025).
73. РосПатент [Электронный ресурс]: Федеральная служба по интеллектуальной собственности. – URL: <https://fips.ru> (дата обращения: 31.05.2025).
74. ВАК [Электронный ресурс] : официальный сайт / Министерство науки и высшего образования Российской Федерации. – URL: <https://vak.minobrnauki.gov.ru> (дата обращения: 31.05.2025).
75. Scopus [Электронный ресурс] : международная реферативная база данных / Elsevier. – URL: <https://www.scopus.com> (дата обращения: 31.05.2025).
76. РИНЦ [Электронный ресурс] : официальный сайт / Издательство «Наука»; поддерживается Российской академией наук. – URL: https://elibrary.ru/projects/citation/cit_index.asp (дата обращения: 31.05.2025).
77. Шутков, М. А. Совершенствование нейронной сети U-Net для сегментации кровеносных сосудов сетчатки [Электронный ресурс] : магистерская диссертация / М. А. Шутков ; Министерство науки и высшего

образования Российской Федерации, Федеральное государственное автономное образовательное учреждение высшего образования "Уральский федеральный университет имени первого Президента России Б.Н. Ельцина". — Екатеринбург, 2024. — 1 электрон. файл (PDF; 1,69 МБ). — Загл. с экрана. — URL: https://elar.urfu.ru/bitstream/10995/140596/1/m_th_m.a.shutkov_2024.pdf (дата обращения: 03.06.2025).

78. Организационная структура клиники [Электронный ресурс] : статья. — URL: <https://primusinpelagus.com/blog/orgstruktura-kliniki/> (дата обращения: 05.04.2025).

79. The Role of Data Lakes in Healthcare [Электронный ресурс] : электронный документ / Perficient. — 2022. — Формат файла: PDF. — Режим доступа: <https://www.perficient.com/-/media/files/guide-pdf-links/the-role-of-data-lakes-in-healthcare.pdf>, свободный (дата обращения: 05.04.2025).

80. Comprehensive eye exams [Электронный ресурс] : официальный сайт American Optometric Association. — URL: <https://www.aoa.org/healthy-eyes/caring-for-your-eyes/eye-exams> (дата обращения: 05.04.2025).

81. Huang S.C., Pareek A., Seyyedi S. et al. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines // *npj Digital Medicine*. — 2020. — Т. 3. — № 136. — DOI: 10.1038/s41746-020-00341-z. — Режим доступа: <https://www.nature.com/articles/s41746-020-00341-z> (дата обращения: 05.04.2025).

82. Zhang, Y., Wang, J., Li, T. Ophthalmology image segmentation using deep learning methods: a review // *Computerized Medical Imaging and Graphics*. — 2021. — Vol. 88. — Article 101845. — DOI: 10.1016/j.compmedimag.2020.101845.

83. Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных» // Система ГАРАНТ. — URL: <https://www.garant.ru/products/ipo/prime/doc/41983> (дата обращения: 01.06.2025)

84. Общий регламент защиты данных (GDPR). Перевод на русский язык с комментариями // Текст GDPR на русском. — URL: <https://gdpr-russia.com> (дата обращения: 01.06.2025)

85. Morgan & Westfield. Business Valuation & Return on Investment (ROI) [Электронный ресурс] — Режим доступа: <https://morganandwestfield.com/knowledge/business-valuation-return-on-investment-roi/> (дата обращения: 01.06.2025).

86. Шевелева О. ЕМИАС — здравоохранение будущего // *Врач и информационные технологии*. 2012. №6. URL: <https://cyberleninka.ru/article/n/emias-zdravoohranenie-buduschego> (дата обращения: 31.05.2025).

87. DRIVE Dataset

88. Цифровой набор изображений сетчатки глаза для выделения сосудов (Digital Retinal Images for Vessel Extraction — DRIVE) [Электронный ресурс] : набор данных для сегментации ретинальных сосудов. — Режим доступа: <https://drive.grand-challenge.org/>, свободный. — Загл. с экрана. — Дата обращения: 01.06.2025

89. STARE Dataset. Structured Analysis of the Retina — STARE [Электронный ресурс] : набор данных для сегментации изображений сетчатки глаза. — Режим доступа: <https://cecas.clemson.edu/~ahoover/stare/> , свободный. — Загл. с экрана. — Дата обращения: 01.06.2025
90. NumPy [Электронный ресурс] : библиотека для научных вычислений на языке Python. — Режим доступа: <https://numpy.org/> , свободный. — Загл. с экрана. — Дата обращения: 01.06.2025
91. OpenCV (Open Source Computer Vision Library) [Электронный ресурс]. — Режим доступа: <https://opencv.org/> , свободный. — Загл. с экрана. — Дата обращения: 01.06.2025
92. Python [Электронный ресурс] : язык программирования. — Режим доступа: <https://www.python.org/> , свободный. — Загл. с экрана. — Дата обращения: 01.06.2025.
93. Keras [Электронный ресурс] : документация и руководства по использованию библиотеки глубокого обучения. — Режим доступа: <https://keras.io/> , свободный. — Загл. с экрана. — Дата обращения: 01.06.2025.
94. Brown, C. D. Receiver operating characteristic curves and related decision measures: a tutorial / Christopher D. Brown, Herbert T. Davis // *Chemometrics and Intelligent Laboratory Systems* . — 2006. — Vol. 80. — P. 24–38. — DOI:10.1016/j.chemolab.2005.05.004.
95. Братко, И. Machine learning: Between accuracy and interpretability / Иван Братко // *Learning, Networks and Statistics* : сб. науч. тр. — Vienna : Springer Vienna, 1997. — P. 163–177. — (International Centre for Mechanical Sciences ; vol. 382). — ISBN нд. — Текст : непосредственный // Научная электронная библиотека. — URL: <https://openalex.org/W1921840746> (дата обращения: 01.06.2025)
96. Сасаки, Ю. The truth of the F-measure / Ютака Сасаки // *Teach Tutor Mater.* — 2007. — Vol. 1, № 5. — P. 1–5.
97. Руби, У., Яндопалли, В. Бинарная кросс-энтропия с использованием техники глубокого обучения для классификации изображений // *Международный журнал передовых тенденций в компьютерных науках и инженерии.* — 2020. — Т. 9, № 10. — С. 100-120
98. Getahun M.N., Rogov O.Y., Dylvov D.V., Somov A., Bouridane A., Hamoudi R. FS-Net: Full Scale Network and Adaptive Threshold for Improving Extraction of Micro-Retinal Vessel Structures [Электронный ресурс] // *arXiv.* – 2023. – *arXiv:2311.08059.* – Режим доступа: <https://arxiv.org/abs/2311.08059>
99. Desiani A., Erwin E., Suprihatin B., Wahyudi Y., Cahyono E., Arhami M. BVU-Net: A U-Net Modification by VGG-Batch Normalization for Retinal Blood Vessel Segmentation // *International Journal of Intelligent Engineering & Systems.* – 2022. – Т. 15, № 4. – С. 303–314. – DOI: 10.22158/ijies.v15n4p303
100. Sanjeevani; Yadav A.K., Akbar M., Kumar M., Yadav D. Retinal blood vessel segmentation using a deep learning method based on modified U-NET model // *Multimedia Tools and Applications.* – 2024. – Т. 83. – С. 82659–82678. – DOI: 10.1007/s11042-024-23220-z

101. Kim H.J., Eesaar H., Chong K.T. Transformer-Enhanced Retinal Vessel Segmentation for Diabetic Retinopathy Detection Using Attention Mechanisms and Multi-Scale Fusion // *Applied Sciences*. – 2024. – T. 14, № 22. – C. 10658. – DOI: 10.3390/app142210658
102. Bhuiya S., Choudhury S.R., Aich G., Maurya M., Sen A. Retinal blood vessel segmentation and analysis using lightweight spatial attention based CNN and data augmentation // *Proceedings of the 2022 IEEE Calcutta Conference (CALCON)*, Kolkata, India, 10–11 December 2022. – P. 122–127. – DOI: 10.1109/CALCON48930.2022.10004389
103. Khan T.M., Alhussein M., Aurangzeb K., Arsalan M., Naqvi S.S., Nawaz S.J. Residual connection-based encoder decoder network (RCED-Net) for retinal vessel segmentation // *IEEE Access*. – 2020. – T. 8. – C. 131257–131272. – DOI: 10.1109/ACCESS.2020.3002978
104. Yin P., Yuan R., Cheng Y., Wu Q. Deep guidance network for biomedical image segmentation // *IEEE Access*. – 2020. – T. 8. – C. 116106–116116. – DOI: 10.1109/ACCESS.2020.2997905
105. Ibtehaz N., Rahman M.S. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation // *Neural Networks*. – 2020. – T. 121. – C. 74–87. – DOI: 10.1016/j.neunet.2019.08.025
106. Yan Z., Yang X., Cheng K.T. Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation // *IEEE Transactions on Biomedical Engineering*. – 2018. – T. 65, № 9. – C. 1912–1923. – DOI: 10.1109/TBME.2017.2784062
107. Cao J., Chen J., Gu Y., Liu J. MFA-UNet: A vessel segmentation method based on multi-scale feature fusion and attention module // *Frontiers in Neuroscience*. – 2023. – T. 17. – C. 1249331. – DOI: 10.3389/fnins.2023.1249331
108. Ye Z., Liu Y., Jing T., He Z., Zhou L. A High-Resolution Network with Strip Attention for Retinal Vessel Segmentation // *Sensors*. – 2023. – T. 23, № 21. – C. 8899. – DOI: 10.3390/s23218899
109. Wu Y., Xia Y., Song Y., Zhang D., Liu D., Zhang C., Cai W. Vessel-Net: Retinal vessel segmentation under multi-path supervision // *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, 13–17 October 2019; Proceedings, Part I 22*. – Berlin/Heidelberg: Springer, 2019. – C. 264–272. – DOI: 10.1007/978-3-030-32248-9_30

ПРИЛОЖЕНИЕ А

```
from tensorflow.keras.models import Model
from tensorflow.keras.layers import (
    Input, Conv2D, MaxPooling2D, Conv2DTranspose, concatenate,
    BatchNormalization, Activation, Reshape, MultiHeadAttention,
    LayerNormalization, Dense, GlobalAveragePooling2D, Add,
    Dropout, Flatten, Resizing, UpSampling2D
)
from tensorflow.keras.applications import ResNet50
import os
import numpy as np
import tensorflow as tf
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.utils import normalize
from tensorflow.keras.callbacks import ModelCheckpoint
# =====
# 1. Базовые компоненты
def conv_block(x, filters, kernel_size=3, use_bn=True):
    x = Conv2D(filters, kernel_size, padding='same')(x)
    if use_bn:
        x = BatchNormalization()(x)
    x = Activation('relu')(x)
    return x

def transformer_block(x, num_heads=4, projection_dim=64):
    # Нормализация и Self-Attention
    x1 = LayerNormalization(epsilon=1e-6)(x)
    attention = MultiHeadAttention(
        num_heads=num_heads,
        key_dim=projection_dim,
        dropout=0.1
    )(x1, x1)

    # Пропускное соединение
    x2 = Add()([attention, x])
    x3 = LayerNormalization(epsilon=1e-6)(x2)

    # Feed Forward Network
    x3 = Dense(projection_dim * 4, activation="gelu")(x3)
    x3 = Dense(projection_dim)(x3)
    return Add()([x3, x2])

# =====
# 2. Архитектуры моделей
# =====
def build_unet(input_shape=(256, 256, 3)):
    inputs = Input(input_shape)

    # Encoder
    c1 = conv_block(inputs, 64)
    c1 = conv_block(c1, 64)
    p1 = MaxPooling2D((2, 2))(c1)

    c2 = conv_block(p1, 128)
    c2 = conv_block(c2, 128)
```

```

p2 = MaxPooling2D((2, 2))(c2)

c3 = conv_block(p2, 256)
c3 = conv_block(c3, 256)
p3 = MaxPooling2D((2, 2))(c3)

c4 = conv_block(p3, 512)
c4 = conv_block(c4, 512)
p4 = MaxPooling2D((2, 2))(c4)

# Bottleneck
b = conv_block(p4, 1024)
b = conv_block(b, 1024)

# Decoder
u1 = Conv2DTranspose(512, (2, 2), strides=(2, 2), padding='same')(b)
u1 = concatenate([u1, c4])
c5 = conv_block(u1, 512)
c5 = conv_block(c5, 512)

u2 = Conv2DTranspose(256, (2, 2), strides=(2, 2), padding='same')(c5)
u2 = concatenate([u2, c3])
c6 = conv_block(u2, 256)
c6 = conv_block(c6, 256)

u3 = Conv2DTranspose(128, (2, 2), strides=(2, 2), padding='same')(c6)
u3 = concatenate([u3, c2])
c7 = conv_block(u3, 128)
c7 = conv_block(c7, 128)

u4 = Conv2DTranspose(64, (2, 2), strides=(2, 2), padding='same')(c7)
u4 = concatenate([u4, c1])
c8 = conv_block(u4, 64)
c8 = conv_block(c8, 64)

# Output
outputs = Conv2D(1, (1, 1), activation='sigmoid')(c8)
return Model(inputs, outputs)

def build_resnet50_segmentation(input_shape=(256, 256, 3)):
    base = ResNet50(
        weights='imagenet',
        include_top=False,
        input_shape=input_shape
    )

    # Замораживаем начальные слои
    for layer in base.layers[:50]:
        layer.trainable = False

    # Skip connections
    s1 = base.get_layer('conv1_relu').output
    s2 = base.get_layer('conv2_block3_out').output
    s3 = base.get_layer('conv3_block4_out').output
    s4 = base.get_layer('conv4_block6_out').output
    bottleneck = base.output

    # Decoder

```

```

    u1 = Conv2DTranspose(512, (2, 2), strides=(2, 2),
padding='same')(bottleneck)
    u1 = concatenate([u1, s4])
    u1 = conv_block(u1, 512)

    u2 = Conv2DTranspose(256, (2, 2), strides=(2, 2), padding='same')(u1)
    u2 = concatenate([u2, s3])
    u2 = conv_block(u2, 256)

    u3 = Conv2DTranspose(128, (2, 2), strides=(2, 2), padding='same')(u2)
    u3 = concatenate([u3, s2])
    u3 = conv_block(u3, 128)

    u4 = Conv2DTranspose(64, (2, 2), strides=(2, 2), padding='same')(u3)
    u4 = concatenate([u4, s1])
    u4 = conv_block(u4, 64)

    # Output
    outputs = Conv2D(1, (1, 1), activation='sigmoid')(u4)
    return Model(base.input, outputs)

def build_ct_unet(input_shape=(256, 256, 3)):
    inputs = Input(input_shape)

    # Encoder (U-Net)
    c1 = conv_block(inputs, 64)
    c1 = conv_block(c1, 64)
    p1 = MaxPooling2D((2, 2))(c1)

    c2 = conv_block(p1, 128)
    c2 = conv_block(c2, 128)
    p2 = MaxPooling2D((2, 2))(c2)

    c3 = conv_block(p2, 256)
    c3 = conv_block(c3, 256)
    p3 = MaxPooling2D((2, 2))(c3)

    c4 = conv_block(p3, 512)
    c4 = conv_block(c4, 512)
    p4 = MaxPooling2D((2, 2))(c4)
    # Bottleneck with Transformer
    b = conv_block(p4, 512)
    b = conv_block(b, 512)

    # Подготовка для трансформера
    _, h, w, c = b.shape
    b_flat = Reshape((h*w, c))(b)
    b_trans = transformer_block(b_flat)
    b_trans = Reshape((h, w, c))(b_trans)

    # Decoder (U-Net)
    u1 = Conv2DTranspose(512, (2, 2), strides=(2, 2), padding='same')(b)
    u1 = concatenate([u1, c4])
    c5 = conv_block(u1, 512)
    c5 = conv_block(c5, 512)

    u2 = Conv2DTranspose(256, (2, 2), strides=(2, 2), padding='same')(c5)
    u2 = concatenate([u2, c3])
    c6 = conv_block(u2, 256)

```

```

c6 = conv_block(c6, 256)

u3 = Conv2DTranspose(128, (2, 2), strides=(2, 2), padding='same')(c6)
u3 = concatenate([u3, c2])
c7 = conv_block(u3, 128)
c7 = conv_block(c7, 128)

u4 = Conv2DTranspose(64, (2, 2), strides=(2, 2), padding='same')(c7)
u4 = concatenate([u4, c1])
c8 = conv_block(u4, 64)
c8 = conv_block(c8, 64)
return Model(inputs, outputs)

def build_unet_resnet_backbone(input_shape=(256, 256, 3)):
    base = ResNet50(
        weights='imagenet',
        include_top=False,
        input_shape=input_shape
    )

    # Skip connections
    s1 = base.get_layer('conv1_relu').output
    s2 = base.get_layer('conv2_block3_out').output
    s3 = base.get_layer('conv3_block4_out').output
    s4 = base.get_layer('conv4_block6_out').output
    bottleneck = base.output

    # Decoder U-Net
    u1 = Conv2DTranspose(512, (2, 2), strides=(2, 2),
padding='same')(bottleneck)
    u1 = concatenate([u1, s4])
    u1 = conv_block(u1, 512)

    u2 = Conv2DTranspose(256, (2, 2), strides=(2, 2), padding='same')(u1)
    u2 = concatenate([u2, s3])
    u2 = conv_block(u2, 256)

    u3 = Conv2DTranspose(128, (2, 2), strides=(2, 2), padding='same')(u2)
    u3 = concatenate([u3, s2])
    u3 = conv_block(u3, 128)

    u4 = Conv2DTranspose(64, (2, 2), strides=(2, 2), padding='same')(u3)
    u4 = concatenate([u4, s1])
    u4 = conv_block(u4, 64)

    outputs = Conv2D(1, (1, 1), activation='sigmoid')(u4)
    return Model(base.input, outputs)

# 3. Мета-модель для стэкинга
def build_meta_model(input_shape=(256, 256, 4)):
    inputs = Input(shape=input_shape) # 4 канала от базовых моделей

    x = Conv2D(64, (3, 3), padding='same', activation='relu')(inputs)
    x = BatchNormalization()(x)
    x = Conv2D(1, (1, 1), activation='sigmoid')(x)

    return Model(inputs, x)

# 4. Функция для построения ансамбля

```

```

def build_ensemble(models, input_shape=(256, 256, 3)):
    # Входные данные
    img_input = Input(shape=input_shape)

    # Прогон через базовые модели
    outputs = [model(img_input) for model in models]

    # Объединение выходов
    stacked = concatenate(outputs, axis=-1)

    # Мета-модель
    final_output = meta_model(stacked)
    return Model(inputs=img_input, outputs=final_output)

# =====

# Инициализация моделей
# =====
# Создаем базовые модели
UNET_MODEL = build_unet()
RESNET_MODEL = build_resnet50_segmentation()
CTUNET_MODEL = build_ct_unet()
UNET_RESNET_MODEL = build_unet_resnet_backbone()

# Создаем мета-модель
META_MODEL = build_meta_model()

# Собираем ансамбль
ENSEMBLE_MODEL = build_ensemble(
    models=[UNET_MODEL, RESNET_MODEL, CTUNET_MODEL, UNET_RESNET_MODEL]
)

# Компиляция и обучение
def train_ensemble(X_train, y_train, X_val, y_val, epochs=50):
    ensemble_model.compile(
        optimizer='adam',
        loss='binary_crossentropy',
        metrics=['accuracy']
    )

    checkpoint = ModelCheckpoint(
        'best_ensemble.h5',
        monitor='val_loss',
        save_best_only=True
    )

    history = ensemble_model.fit(
        X_train, y_train,
        validation_data=(X_val, y_val),
        epochs=epochs,
        batch_size=8,
        callbacks=[checkpoint]
    )

    return history

def load_dataset(data_dir, img_size=(256, 256)):
    image_paths = []
    mask_paths = []

```

```

# DRIVE
drive_img_dir = os.path.join(data_dir, 'DRIVE', 'training', 'images')
drive_mask_dir = os.path.join(data_dir, 'DRIVE', 'training', 'masks')
for fname in os.listdir(drive_img_dir):
    if fname.endswith('.tif'):
        base_name = fname.split('_')[0]
        image_paths.append(os.path.join(drive_img_dir, fname))
        mask_paths.append(os.path.join(drive_mask_dir,
f"{base_name}_manuall.gif"))

# STARE
stare_img_dir = os.path.join(data_dir, 'STARE', 'images')
stare_mask_dir = os.path.join(data_dir, 'STARE', 'labels')
for fname in os.listdir(stare_img_dir):
    if fname.endswith('.ppm'):
        base_name = fname.split('.')[0]
        image_paths.append(os.path.join(stare_img_dir, fname))
        mask_paths.append(os.path.join(stare_mask_dir, f"{base_name}.ppm"))

return image_paths, mask_paths

def preprocess_data(image_paths, mask_paths, img_size=(256, 256)):
    images = []
    masks = []

    for img_path, mask_path in zip(image_paths, mask_paths):
        # Load and resize image
        img = tf.keras.preprocessing.image.load_img(img_path,
target_size=img_size, color_mode='rgb')
        img = tf.keras.preprocessing.image.img_to_array(img)
        img = img / 255.0

        # Load and resize mask
        mask = tf.keras.preprocessing.image.load_img(mask_path,
target_size=img_size, color_mode='grayscale')
        mask = tf.keras.preprocessing.image.img_to_array(mask)
        mask[mask > 0] = 1.0 # binary mask

        images.append(img)
        masks.append(mask)

    return np.array(images), np.array(masks)

DATASET_PATH = './dataset/' # Укажи свой путь

image_paths, mask_paths = load_dataset(DATASET_PATH)
X, y = preprocess_data(image_paths, mask_paths)

# Разделение на train/val
from sklearn.model_selection import train_test_split
X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.2,
random_state=42)

print("Train shape:", X_train.shape, y_train.shape)
print("Val shape:", X_val.shape, y_val.shape)
history = train_ensemble(X_train, y_train, X_val, y_val, epochs=50)
ensemble_model.save('final_ensemble_model.h5')

```

