Министерство образования Республики Беларусь

Учреждение образования

«Брестский государственный технический университет»

Факультет электронно-информационных систем

Кафедра интеллектуальных информационных технологий

Отчет

о прохождении преддипломной практики

в OOO «Аира стиль»

Студента факультета электронно-информационных систем

4 курс, группа ИИ-21 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Т.В. Литвинюк

(подпись)

Руководитель практики

от университета

Старший Преподаватель \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Н.С. Монтик

(подпись)

Руководитель практики

от предприятия

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Бартош Е. В.

(подпись, печать)

Брест 2025

**СОДЕРЖАНИЕ**

[ВВЕДЕНИЕ 3](#_Toc193624344)

[1. СИСТЕМНЫЙ АНАЛИЗ И ПОСТАНОВКА ЗАДАЧИ 5](#_Toc193624345)

[1.1 Описание предметной области 5](#_Toc193624346)

[1.2 Анализ существующих решений 6](#_Toc193624347)

[1.3 Основы структуры человеческого глаза 7](#_Toc193624348)

[1.4 Основы тонометрии 8](#_Toc193624349)

[1.5 Обоснование необходимости реализации 9](#_Toc193624350)

[1.6 Постановка задачи 10](#_Toc193624351)

[2. ПРОЕКТИРОВАНИЕ АЛГОРИТМА 12](#_Toc193624352)

[2.1 Анализ инструментов для реализации 12](#_Toc193624353)

[2.2 Обзор набора данных тонометрии 13](#_Toc193624354)

[2.3 Проектирование архитектуры нейронных сетей 14](#_Toc193624355)

[ЗАКЛЮЧЕНИЕ 17](#_Toc193624356)

[СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ 19](#_Toc193624357)

ВВЕДЕНИЕ

Глаукома представляет собой одну из ведущих причин необратимой слепоты во всем мире, что делает своевременную диагностику и лечение этого заболевания критически важными задачами современной медицины. Это хроническое заболевание глаз характеризуется повышенным внутриглазным давлением, которое со временем может привести к повреждению зрительного нерва и потере зрения. Несмотря на значительный прогресс в области офтальмологии, ранняя диагностика глаукомы остается сложной задачей, поскольку на начальных стадиях она часто протекает бессимптомно. Традиционные методы диагностики, такие как тонометрия, позволяют измерять внутриглазное давление, однако их точность и способность выявлять заболевание на ранних стадиях зависят от опыта врача и качества интерпретации данных.

С развитием технологий искусственного интеллекта и глубокого обучения открываются новые возможности для автоматизации и повышения точности диагностических процессов. Использование нейронных сетей для анализа медицинских изображений и данных тонометрии демонстрирует многообещающие результаты в выявлении патологий глаза, включая глаукому. Такие подходы позволяют не только ускорить процесс диагностики, но и снизить вероятность человеческой ошибки, что особенно важно в условиях ограниченного доступа к квалифицированным специалистам. В данном контексте разработка системы классификации глаукомы на основе глубокого обучения становится актуальной задачей, способной внести значительный вклад в улучшение качества медицинской помощи.

Целью данной дипломной работы является разработка и реализация модели глубокого обучения для классификации глаукомы на основе данных тонометрии глаза с использованием доступного набора данных. В работе будет рассмотрена предметная область, проведен анализ существующих решений, а также спроектирована и обучена нейронная сеть, способная эффективно различать здоровые глаза от глаз с признаками глаукомы. Особое внимание будет уделено выбору подходящих инструментов, обработке данных и оптимизации модели для достижения высокой точности классификации.

Предложенная разработка основывается на наборе данных "Glaucoma Dataset EyePACS AIROGS Light v2", доступном на платформе Kaggle. Этот набор включает изображения глазного дна, которые будут использованы для обучения и тестирования модели. В процессе работы будут применены современные методы обработки изображений и архитектуры нейронных сетей, такие как MobileNet V3, что позволит создать эффективное решение, пригодное для практического применения. Итогом исследования станет обученная модель, сохраненная для дальнейшего использования, а также анализ ее производительности на тестовых данных.

Таким образом, данная работа направлена на решение актуальной проблемы ранней диагностики глаукомы с использованием передовых технологий машинного обучения. Результаты исследования могут послужить основой для создания автоматизированных систем поддержки принятия решений в офтальмологии, что в перспективе улучшит доступность и качество диагностики этого серьезного заболевания. В следующих разделах будут подробно рассмотрены теоретические основы, методология разработки и практическая реализация поставленной задачи.

1. СИСТЕМНЫЙ АНАЛИЗ И ПОСТАНОВКА ЗАДАЧИ

# Описание предметной области

Предметная область данной дипломной работы охватывает сферу офтальмологии, а именно диагностику глаукомы — одного из наиболее распространенных и опасных заболеваний глаз, которое может привести к необратимой потере зрения. Глаукома представляет собой группу хронических заболеваний, основным признаком которых является повышение внутриглазного давления (ВГД), вызывающее постепенное повреждение зрительного нерва. Этот процесс, как правило, протекает медленно и на ранних стадиях не сопровождается явными симптомами, что значительно затрудняет своевременное выявление болезни. По данным Всемирной организации здравоохранения, глаукома занимает второе место среди причин слепоты в мире, уступая лишь катаракте, и затрагивает миллионы людей, особенно пожилого возраста. Таким образом, разработка методов ранней диагностики этого заболевания является одной из приоритетных задач современной медицины.

В основе диагностики глаукомы лежит измерение внутриглазного давления с помощью тонометрии — метода, который позволяет оценить давление жидкости внутри глазного яблока. Тонометрия может проводиться различными способами, включая контактные и бесконтактные техники, и часто дополняется анализом изображений глазного дна, где можно наблюдать изменения зрительного нерва и сосудистой системы. Однако традиционные подходы к интерпретации данных тонометрии и визуального осмотра имеют свои ограничения: они требуют высокой квалификации врача, а также могут быть подвержены субъективным ошибкам. Кроме того, доступ к специализированным офтальмологическим услугам ограничен в некоторых регионах, что подчеркивает необходимость разработки автоматизированных и доступных решений.

С развитием технологий искусственного интеллекта (ИИ) и глубокого обучения (Deep Learning, DL) в медицинской практике появляется возможность значительно улучшить диагностические процессы. Глубокое обучение, основанное на использовании многослойных нейронных сетей, демонстрирует выдающиеся результаты в задачах анализа изображений и классификации данных. В контексте офтальмологии такие технологии уже успешно применяются для выявления таких заболеваний, как диабетическая ретинопатия и возрастная макулярная дегенерация. Применение аналогичных методов для диагностики глаукомы открывает перспективы для создания систем, которые смогут автоматически анализировать данные тонометрии и изображения глазного дна, предоставляя врачам надежные инструменты для принятия решений.

Предметная область данной работы также включает изучение особенностей наборов данных, используемых для обучения моделей глубокого обучения. В частности, в исследовании будет задействован набор данных "Glaucoma Dataset EyePACS AIROGS Light v2", содержащий изображения глазного дна с разметкой, указывающей наличие или отсутствие признаков глаукомы. Эти данные представляют собой ценный ресурс для разработки и тестирования алгоритмов классификации, поскольку они отражают реальные клинические случаи. Важным аспектом является также предварительная обработка данных, включая аугментацию изображений и нормализацию, что позволяет повысить устойчивость модели к вариациям в качестве изображений и условиям их получения.

# Анализ существующих решений

На сегодняшний день автоматизированные методы диагностики глаукомы активно развиваются, и существует несколько подходов к обработке медицинских изображений и данных тонометрии. Основные направления исследований включают традиционные методы обработки изображений, машинное обучение и глубокие нейронные сети.

**Традиционные методы анализа**

Классические методы диагностики глаукомы основаны на анализе изображений глазного дна и параметров внутриглазного давления. Для этого применяются различные алгоритмы обработки изображений, такие как:

* Выделение контуров и анализ формы зрительного нерва;
* Сегментация слоя нервных волокон сетчатки;
* Оценка толщины роговицы и ее влияния на показания тонометрии.

Эти методы позволяют выявлять патологические изменения, но требуют значительных вычислительных ресурсов и высокой квалификации врача для интерпретации результатов.

**Методы машинного обучения**

В последние годы получили развитие алгоритмы машинного обучения, такие как:

* SVM (метод опорных векторов);
* KNN (метод ближайших соседей);
* Решающие деревья и случайные леса.

Эти алгоритмы используют вручную отобранные признаки изображений глазного дна и параметры тонометрии для обучения модели, которая затем классифицирует пациентов на группы с высоким и низким риском глаукомы.

**Глубокое обучение**

С появлением глубоких нейронных сетей значительно улучшились результаты автоматической диагностики. Используются модели, такие как:

* CNN (сверточные нейронные сети) для анализа изображений глазного дна;
* RNN и трансформеры для обработки последовательностей данных;
* Гибридные модели, комбинирующие изображения и числовые данные тонометрии.

В частности, архитектуры ResNet, MobileNet и EfficientNet показывают высокую точность в задачах классификации глаукомы. Они способны автоматически выделять важные признаки из изображений и анализировать их с минимальным вмешательством человека.

**Сравнительный анализ**

Исследования показывают, что глубокие нейронные сети превосходят традиционные методы и алгоритмы машинного обучения по точности диагностики глаукомы. Однако их применение требует больших объемов размеченных данных и вычислительных мощностей. Кроме того, важными факторами остаются интерпретируемость моделей и их клиническая применимость.

Таким образом, использование глубокого обучения для автоматической классификации глаукомы является перспективным направлением, позволяющим повысить точность диагностики и снизить зависимость от субъективного фактора врача.

# Основы структуры человеческого глаза

Понимание структуры человеческого глаза является фундаментальной основой для разработки методов диагностики глаукомы, поскольку это заболевание непосредственно связано с анатомическими и физиологическими особенностями органа зрения. Человеческий глаз представляет собой сложную оптическую систему, которая обеспечивает восприятие света, его преобразование в нервные импульсы и передачу информации в мозг. Знание строения глаза позволяет лучше интерпретировать данные, получаемые в ходе тонометрии и анализа изображений глазного дна, а также понять механизмы развития глаукомы. В данном разделе будут рассмотрены ключевые элементы структуры глаза, их функции и связь с патологическими процессами, характерными для этого заболевания.

Человеческий глаз имеет форму шара диаметром около 24–25 мм и состоит из нескольких основных структурных компонентов. Внешний слой глаза, называемый склерой, представляет собой плотную соединительную ткань, которая защищает внутренние элементы и поддерживает форму органа. Передняя часть склеры переходит в прозрачную роговицу, которая выполняет функцию первичной линзы, преломляющей световые лучи. Роговица играет важную роль в измерении внутриглазного давления, поскольку ее толщина и кривизна могут влиять на результаты тонометрии. За роговицей расположена передняя камера глаза, заполненная водянистой влагой — прозрачной жидкостью, которая поддерживает давление внутри глаза и питает окружающие ткани.

Следующий важный элемент — радужка, которая регулирует количество света, поступающего в глаз, за счет изменения размера зрачка. За радужкой находится хрусталик — прозрачное тело, способное менять свою форму для фокусировки изображения на сетчатке. Водянистая влага, циркулирующая между передней и задней камерами глаза, оттекает через трабекулярную сеть — структуру, расположенную в углу передней камеры. Нарушение оттока этой жидкости является одной из основных причин повышения внутриглазного давления при глаукоме, что делает трабекулярную сеть ключевым объектом изучения в контексте данного заболевания. Повышенное давление оказывает воздействие на заднюю часть глаза, где расположены сетчатка и зрительный нерв.

Сетчатка представляет собой внутренний слой глаза, состоящий из фоторецепторных клеток (палочек и колбочек), которые преобразуют свет в электрические сигналы. Эти сигналы передаются через зрительный нерв в мозг для дальнейшей обработки. Зрительный нерв выходит из глаза через диск зрительного нерва — область, видимую на снимках глазного дна. При глаукоме повышенное внутриглазное давление может сдавливать и повреждать волокна зрительного нерва, что приводит к его атрофии и постепенной утрате поля зрения. Именно изменения в структуре диска зрительного нерва и слоя нервных волокон сетчатки являются важными диагностическими признаками, которые можно выявить с помощью изображений глазного дна.

Глазное яблоко также окружено стекловидным телом — гелеобразной субстанцией, которая заполняет пространство между хрусталиком и сетчаткой, обеспечивая поддержку формы глаза. Хотя стекловидное тело непосредственно не связано с развитием глаукомы, его состояние может влиять на качество изображений, используемых для диагностики. Кроме того, кровеносные сосуды, питающие сетчатку, также видны на снимках глазного дна и могут служить дополнительным индикатором патологических изменений, таких как сужение или аномалии, связанные с глаукомой.

# Основы тонометрии

**Тонометрия** – это метод измерения внутриглазного давления (ВГД), который играет ключевую роль в диагностике и мониторинге глаукомы. ВГД определяется балансом между выработкой и оттоком внутриглазной жидкости и является важным показателем здоровья глаз.

**Нормальные значения внутриглазного давления**

Среднее нормальное значение ВГД составляет **10–21 мм рт. ст.** Отклонения от этих значений могут указывать на развитие глаукомы или других офтальмологических заболеваний.

**Методы тонометрии**

1. **Апланационная тонометрия (по Гольдману)**
   * Является «золотым стандартом» измерения ВГД.
   * Измерение основано на приложении давления к роговице для оценки силы, необходимой для ее уплощения.
   * Требует местного обезболивания и использования флуоресцентного красителя.
2. **Бесконтактная (пневматическая) тонометрия**
   * Использует воздушный поток для сплющивания роговицы.
   * Не требует обезболивания, но может быть менее точной по сравнению с апланационной тонометрией.
3. **Динамическая контурная тонометрия**
   * Учитывает индивидуальные характеристики роговицы, такие как ее толщина и эластичность.
   * Позволяет получить более точные измерения при измененных свойствах роговицы.
4. **Импрессионная тонометрия (по Маклакову)**
   * Основана на измерении глубины вдавливания специального грузика в роговицу.
   * Используется преимущественно в странах СНГ и требует местного обезболивания.
5. **Электронная тонометрия**
   * Современные портативные электронные тонометры обеспечивают быструю и точную оценку ВГД.

**Факторы, влияющие на точность измерений**

* **Толщина роговицы** – более толстая роговица может завышать показатели ВГД, а более тонкая – занижать.
* **Время суток** – ВГД может колебаться в течение дня, достигая пика утром.
* **Положение тела** – в лежачем положении ВГД обычно выше, чем в сидячем.

Тонометрия является одним из важнейших методов в офтальмологии и широко применяется для раннего выявления глаукомы. Однако для более точной диагностики ее результаты необходимо анализировать в совокупности с другими методами исследования.

# Обоснование необходимости реализации

Глаукома является одной из ведущих причин необратимой слепоты во всем мире. По данным Всемирной организации здравоохранения (ВОЗ), миллионы людей ежегодно теряют зрение из-за поздней диагностики этого заболевания. Основной причиной столь высокой заболеваемости и осложненного выявления глаукомы является бессимптомное течение на ранних стадиях, что затрудняет своевременное обнаружение патологии.

Одним из ключевых методов диагностики глаукомы является **тонометрия**, позволяющая измерять внутриглазное давление (ВГД) и выявлять отклонения от нормы. Однако существующие методы диагностики требуют высокой квалификации врача и могут зависеть от субъективных факторов, таких как интерпретация результатов офтальмологом, состояние пациента и технические особенности используемого оборудования. Это приводит к необходимости создания более точных, автоматизированных и доступных методов диагностики.

В последние годы развитие технологий искусственного интеллекта, в частности **глубокого обучения**, открыло новые возможности для анализа медицинских данных. Глубокие нейронные сети уже доказали свою эффективность в таких областях медицины, как радиология, дерматология и офтальмология. Их способность анализировать сложные паттерны в изображениях делает их перспективным инструментом для диагностики глаукомы на основе данных тонометрии.

Разработка автоматизированной модели диагностики глаукомы на основе глубокого обучения необходима по следующим причинам:

1. **Повышение точности диагностики**
   * Исключение субъективного человеческого фактора при интерпретации данных.
   * Использование современных архитектур нейросетей, способных выявлять закономерности, незаметные для человека.
2. **Ранняя диагностика заболевания**
   * Глубокие нейронные сети могут находить малозаметные признаки глаукомы, позволяя выявлять болезнь на ранних стадиях.
   * Своевременное лечение может значительно замедлить прогрессирование заболевания и сохранить зрение.
3. **Автоматизация и доступность диагностики**
   * Разработка модели, способной анализировать изображения глаз и показатели ВГД без необходимости присутствия офтальмолога.
   * Возможность интеграции технологии в медицинские центры и удаленные диагностические системы.
4. **Оптимизация временных и материальных затрат**
   * Уменьшение нагрузки на медицинский персонал за счет автоматического анализа изображений.
   * Снижение затрат на диагностику благодаря использованию программного анализа вместо дорогостоящих офтальмологических обследований.
5. **Развитие цифровой медицины**
   * Современные технологии позволяют объединять медицинские данные в облачные сервисы, что способствует лучшему мониторингу состояния пациентов.
   * Возможность интеграции разработанной модели в телемедицинские платформы для удаленной диагностики.

Таким образом, реализация модели глубокого обучения для автоматического анализа данных тонометрии глаза является важным и актуальным направлением исследования. Данный проект имеет потенциал значительно повысить точность и доступность диагностики глаукомы, что позволит своевременно выявлять заболевание и улучшить качество жизни пациентов.

# Постановка задачи

Основной задачей данной работы является разработка и обучение модели глубокого обучения для автоматической классификации глаукомы на основе данных тонометрии глаза. Для решения этой задачи необходимо выполнить несколько ключевых этапов, которые включают сбор, предобработку и анализ данных, а также разработку и обучение нейросетевой модели. Важными аспектами являются:

1. **Предобработка и анализ данных**  
   Необходимо собрать и подготовить набор данных, включающий изображения глаз, полученные с использованием тонометрии, а также соответствующие метки, указывающие на наличие или отсутствие глаукомы. В процессе работы с данными нужно выполнить следующие шаги:
   * Провести очистку данных от шума и артефактов.
   * Применить методы аугментации для увеличения объема обучающих данных и повышения обобщающих способностей модели.
   * Подготовить данные для нормализации и стандартизации, чтобы улучшить обучение модели.
2. **Выбор и разработка архитектуры нейронной сети**  
   Необходимо выбрать подходящую архитектуру нейронной сети, способную эффективно классифицировать изображения. Для данной задачи предполагается использование современных моделей глубокого обучения, таких как сверточные нейронные сети (CNN), которые зарекомендовали себя в области обработки изображений. Важно, чтобы выбранная модель могла:
   * Извлекать признаки из изображений глаз, связанные с особенностями развития глаукомы.
   * Обрабатывать изображения разного качества и с различными уровнями шума.
   * Добиться высокой точности классификации при ограниченных вычислительных ресурсах.
3. **Обучение и настройка модели**  
   После выбора архитектуры сети необходимо обучить модель на подготовленных данных. В этом процессе ключевыми задачами являются:
   * Разработка оптимальной стратегии обучения, включая настройку гиперпараметров (например, скорости обучения, размера пакета и числа эпох).
   * Применение методов регуляризации, таких как dropout, для предотвращения переобучения.
   * Оценка производительности модели на валидационном наборе данных и корректировка параметров для достижения наилучших результатов.
4. **Оценка точности и эффективности модели**  
   После обучения модели необходимо провести оценку её точности на тестовом наборе данных. Это включает в себя:
   * Вычисление точности, полноты, F1-меры и других метрик для оценки качества классификации.
   * Анализ ошибок модели и возможных причин недочетов.
   * Сравнение полученных результатов с существующими методами диагностики глаукомы, чтобы выявить преимущества и недостатки разработанного подхода.
5. **Интерпретация и применение результатов**  
   На основе полученной модели необходимо сформулировать рекомендации по её применению в медицинской практике. Это может включать:
   * Разработку программного интерфейса для использования модели в реальных клинических условиях.
   * Оценку возможности интеграции модели в системы телемедицины для удаленной диагностики.
   * Рекомендации по дальнейшему улучшению модели, включая использование дополнительных данных и улучшение архитектуры сети.

Таким образом, постановка задачи заключается в создании эффективного инструмента для диагностики глаукомы на основе анализа данных тонометрии, что позволит улучшить доступность и точность диагностики этого заболевания, а также снизить зависимость от человеческого фактора в интерпретации результатов.

1. ПРОЕКТИРОВАНИЕ АЛГОРИТМА

# 2.1 Анализ инструментов для реализации

Для реализации задачи классификации изображений с использованием глубоких нейронных сетей были выбраны следующие инструменты и библиотеки:

1. **PyTorch** – это мощная библиотека для построения нейронных сетей, которая является основой для разработки и обучения моделей глубокого обучения. PyTorch предоставляет гибкие инструменты для создания, обучения и оптимизации нейросетевых моделей. В данном коде используются основные модули PyTorch, такие как:
   * torch: для работы с тензорами и математическими операциями.
   * torchvision: для работы с изображениями и стандартными предобученными моделями, такими как MobileNet.
   * torch.optim: для оптимизации модели, включая использование различных методов оптимизации (например, Adam).
   * torch.nn: для создания слоев и моделей нейронных сетей.
   * torch.utils.data: для работы с загрузкой и аугментацией данных.
2. **MobileNetV3** – это предварительно обученная модель, предоставленная библиотекой torchvision. Она используется в коде для создания основы модели классификации, благодаря своей легкости и хорошей производительности на мобильных устройствах и других устройствах с ограниченными вычислительными мощностями. В коде заменяется последний слой сети на новый с двумя выходами, что позволяет адаптировать модель для решения задачи бинарной классификации (наличие или отсутствие глаукомы).
3. **Аугментация данных** – для улучшения обучаемости модели используется кастомная аугментация данных, в частности:
   * **Добавление шума** (AddNoise) для увеличения устойчивости модели к шумам.
   * **Цветовая аугментация** с использованием ColorJitter, чтобы модель могла быть более устойчивой к различным цветовым условиям на изображениях.
   * **Геометрическая аугментация** с помощью операций случайного вращения, афинных преобразований и случайных переворотов, что помогает повысить обобщающую способность модели и избегать переобучения.
4. **Обработка данных** – используется стандартная практика предобработки изображений:
   * Изменение размера изображений до 512x512 пикселей с помощью transforms.Resize.
   * Нормализация данных, с использованием заранее рассчитанных средних и стандартных отклонений для изображений, таких как те, что использовались при обучении предварительно обученной модели MobileNetV3.
5. **Использование прогресса в обучении** – в коде используется tqdm, чтобы визуализировать процесс обучения с прогрессом, что помогает отслеживать процесс на каждом этапе обучения и проверки.
6. **Оптимизация** – используется оптимизатор Adam для обновления весов модели. Это популярный метод оптимизации для задач глубокого обучения, который автоматически корректирует шаги обновления весов для каждого параметра, улучшая обучение.
7. **Обучение и тестирование модели** – обучение модели производится с использованием двух фаз:
   * **Обучение (train)** – в процессе обучения модель обновляет свои веса.
   * **Проверка (validation)** – на этой фазе модель оценивает свою работу на валидационном наборе, и происходит выбор наилучшей модели на основе точности.
8. **Параметры обучения**:
   * Модель обучается в течение 6 эпох с использованием функции потерь CrossEntropyLoss, подходящей для задач классификации.
   * Для корректировки скорости обучения используется StepLR scheduler, который уменьшает скорость обучения на 10% каждые 3 эпохи.
9. **Тестирование модели** – после завершения обучения, модель тестируется на отдельном тестовом наборе данных, чтобы оценить её эффективность на невиданных данных.

**Преимущества выбранных инструментов и подходов:**

* **PyTorch** и **MobileNetV3** обеспечивают высокую гибкость и скорость разработки.
* Использование предварительно обученной модели **MobileNetV3** позволяет значительно сократить время на обучение и достичь хороших результатов с минимальными затратами.
* Аугментация данных помогает избежать переобучения и повышает обобщающую способность модели.
* Применение эффективных методов оптимизации (например, Adam) и регуляризации (например, scheduler для LR) позволяет добиться хороших результатов с минимальными вычислительными затратами.

# 2.2 Обзор набора данных тонометрии

Датасет EyePACS-AIROGS-light-v2 представляет собой улучшенную и стандартизированную версию набора изображений глазного дна, предназначенную для обучения моделей машинного обучения для диагностики глаукомы. Он основан на подмножестве изображений из оригинального набора Rotterdam EyePACS AIROGS и включает следующие особенности:

* **Размер изображений**: увеличены с 256x256 до 512x512 пикселей.
* **Формат изображений**: изменен с JPG на PNG для повышения качества.
* **Размеры выборок**:
  + Обучающая выборка: 4000 изображений (~84%).
  + Валидационная выборка: 385 изображений (~8%).
  + Тестовая выборка: 385 изображений (~8%).
* **Классы изображений**: разделены на две категории: referable glaucoma (RG) и non-referable glaucoma (NRG).

Для обеспечения высокого качества разметки, изображения были оценены группой из 20 специалистов, включая офтальмологов и оптометристов, с использованием веб-системы аннотирования. Несогласованные оценки были пересмотрены опытными специалистами для достижения точности диагностики.

Данный датасет способствует разработке и оценке алгоритмов для автоматического обнаружения глаукомы, обеспечивая стандартизированные и высококачественные данные для обучения моделей машинного обучения.

# 2.3 Проектирование архитектуры нейронных сетей

Проектирование архитектуры нейронных сетей является ключевым этапом в разработке системы классификации глаукомы на основе данных тонометрии и изображений глазного дна. Этот процесс включает выбор подходящей модели, настройку ее параметров, определение методов предобработки данных и стратегии обучения, чтобы обеспечить высокую точность и устойчивость системы к реальным условиям. В рамках данной дипломной работы была выбрана архитектура MobileNet V3 Large, которая адаптирована для задачи бинарной классификации (глаукома/нет глаукомы) с использованием набора данных "Glaucoma Dataset EyePACS AIROGS Light v2". В этом разделе подробно описывается процесс проектирования, включая обоснование выбора модели, этапы подготовки данных, модификацию архитектуры и реализацию процесса обучения, а также анализ факторов, влияющих на производительность системы.

**Обоснование выбора архитектуры**

Выбор архитектуры нейронной сети для классификации глаукомы основывался на нескольких критериях: вычислительная эффективность, способность к обобщению, а также успешное применение в задачах анализа медицинских изображений. MobileNet V3 Large, разработанная компанией Google, представляет собой сверточную нейронную сеть (CNN), оптимизированную для работы на устройствах с ограниченными вычислительными ресурсами, таких как мобильные устройства, при сохранении высокой точности классификации. Эта модель сочетает в себе передовые подходы, такие как инвертированные остаточные блоки (Inverted Residuals) и механизм внимания (Squeeze-and-Excitation), что делает ее подходящей для обработки сложных изображений глазного дна. По сравнению с более тяжелыми архитектурами, такими как ResNet или EfficientNet, MobileNet V3 обеспечивает баланс между скоростью обработки и качеством предсказаний, что важно для потенциального практического применения в медицинских системах.

Кроме того, MobileNet V3 изначально обучена на большом наборе данных ImageNet, что позволяет использовать предварительно обученные веса (transfer learning) для ускорения процесса обучения и повышения точности на специализированных задачах, таких как диагностика глаукомы. Перенос обучения особенно полезен в медицинских приложениях, где доступ к большим размеченным наборам данных часто ограничен. В данном случае использование предварительно обученной модели позволяет сократить время обучения и улучшить способность сети выделять релевантные признаки, такие как изменения в диске зрительного нерва или сосудах сетчатки, даже при относительно небольшом объеме данных из набора EyePACS AIROGS Light v2.

**Подготовка данных и аугментация**

Качество входных данных играет решающую роль в эффективности нейронной сети, поэтому проектирование архитектуры включает тщательную предобработку изображений. Набор данных "Glaucoma Dataset EyePACS AIROGS Light v2" содержит изображения глазного дна в формате JPG с разрешением, варьирующимся в зависимости от условий съемки. Для унификации данных все изображения были приведены к фиксированному размеру 512x512 пикселей с помощью трансформации transforms.Resize. Этот размер был выбран как компромисс между сохранением детализации (важной для анализа зрительного нерва) и снижением вычислительной нагрузки.

Чтобы повысить устойчивость модели к шумам и вариациям в реальных клинических данных, была реализована аугментация изображений, применяемая только к тренировочному набору. В частности, использовались следующие преобразования:

* **Добавление шума** (AddNoise с уровнем 0.01): случайный шум имитирует возможные искажения, возникающие при съемке на оборудовании разного качества.
* **Цветовые искажения** (transforms.ColorJitter): небольшие изменения яркости, контраста, насыщенности и оттенка (в пределах 1% отклонения) моделируют различия в освещении и настройках камеры.
* **Горизонтальные и вертикальные отражения** (RandomHorizontalFlip, RandomVerticalFlip): эти преобразования учитывают симметрию глазного дна и увеличивают разнообразие данных.
* **Нормализация** (transforms.Normalize с параметрами mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]): приведение значений пикселей к стандартному распределению, совместимому с предварительно обученными весами MobileNet V3.

Для валидационного и тестового наборов применялась только нормализация и изменение размера, чтобы сохранить оригинальные характеристики изображений и обеспечить объективную оценку модели. Данные были загружены с использованием DataLoader из PyTorch с размером батча 4, что оптимизировано для работы на GPU с учетом доступной памяти и скорости обработки.

**Модификация архитектуры**

Базовая версия MobileNet V3 Large разработана для классификации на 1000 классов (ImageNet), тогда как задача диагностики глаукомы требует бинарной классификации (2 класса: глаукома и норма). Для адаптации модели был изменен последний слой классификатора. Исходный линейный слой с 1280 входами и 1000 выходами был заменен на nn.Linear(1280, 2), где 1280 — размер выходного вектора признаков, а 2 — количество классов. Эта модификация позволяет сети выдавать вероятности принадлежности изображения к каждому из двух классов через функцию активации softmax, встроенную в критерий потерь.

Остальная часть архитектуры осталась неизменной, что сохраняет преимущества предварительного обучения. Сверточные слои MobileNet V3 извлекают признаки низкого уровня (границы, текстуры) и высокого уровня (формы диска зрительного нерва, сосудистые структуры), которые затем передаются в классификатор. Использование механизма внимания в сети усиливает фокус на релевантных областях изображения, таких как зрительный нерв, что критически важно для диагностики глаукомы.

**Процесс обучения и оптимизация**

Обучение модели реализовано с использованием фреймворка PyTorch, что обеспечивает гибкость и поддержку работы на GPU. Процесс включает несколько ключевых компонентов:

* **Критерий потерь**: Для задачи бинарной классификации выбрана функция кросс-энтропии (nn.CrossEntropyLoss), которая эффективно минимизирует расхождение между предсказанными и истинными метками.
* **Оптимизатор**: Использован алгоритм Adam (optim.Adam) с начальной скоростью обучения 0.001. Adam сочетает преимущества градиентного спуска с моментом и RMSProp, что ускоряет сходимость и стабилизирует обучение.
* **Планировщик скорости обучения**: Применен lr\_scheduler.StepLR с шагом 3 эпохи и коэффициентом уменьшения 0.1. Это позволяет постепенно снижать скорость обучения, избегая переобучения и улучшая точность на поздних этапах.
* **Количество эпох**: Установлено 6 эпох, что является компромиссом между временем обучения и достижением приемлемой точности на валидационной выборке.

Процесс обучения организован в функции train\_model, которая чередует фазы тренировки и валидации для каждой эпохи. Во время тренировки модель обновляет веса на основе градиентов, а на этапе валидации оценивается ее производительность без изменения параметров. Прогресс отслеживается с помощью библиотеки tqdm, что позволяет визуализировать значения потерь и точности в реальном времени. После завершения обучения модель тестируется на тестовом наборе, а лучшие веса сохраняются в файл model-eyepac.pth для последующего использования.

**Анализ производительности и потенциальные улучшения**

Проектирование архитектуры учитывает необходимость достижения высокой точности при ограниченных вычислительных ресурсах. MobileNet V3 Large содержит около 5.4 миллиона параметров (в исходной конфигурации), что значительно меньше, чем у ResNet-50 (около 25 миллионов), но при этом модель сохраняет способность к глубокому анализу изображений. Ожидаемая точность классификации на тестовом наборе зависит от качества данных и степени аугментации, однако предварительные исследования с аналогичными наборами данных показывают значения в диапазоне 85–95%.

Для дальнейшего улучшения архитектуры можно рассмотреть следующие подходы:

* Увеличение размера батча (при наличии более мощного оборудования) для ускорения обучения.
* Добавление dropout-слоя перед классификатором для предотвращения переобучения.
* Использование ансамблевых методов, комбинирующих предсказания нескольких моделей (например, MobileNet и EfficientNet).
* Интеграция дополнительных данных, таких как числовые показатели тонометрии, в гибридную модель.

ЗАКЛЮЧЕНИЕ

Разработка системы классификации глаукомы на основе глубокого обучения, представленная в данной дипломной работе, направлена на решение одной из актуальных задач современной офтальмологии — повышение точности и доступности ранней диагностики этого серьезного заболевания. Глаукома, будучи одной из главных причин необратимой слепоты, требует эффективных инструментов для своевременного выявления, особенно на бессимптомных стадиях. Использование технологий искусственного интеллекта и, в частности, глубокого обучения открывает новые перспективы для автоматизации диагностических процессов, что позволяет снизить зависимость от субъективного опыта врача и расширить доступ к качественной медицинской помощи в условиях ограниченных ресурсов.

В ходе исследования была успешно реализована модель на основе архитектуры MobileNet V3 Large, адаптированной для задачи бинарной классификации глаукомы по изображениям глазного дна из набора данных "Glaucoma Dataset EyePACS AIROGS Light v2". Этот набор данных, содержащий размеченные снимки, стал основой для обучения и тестирования модели, а тщательная предобработка, включая аугментацию и нормализацию, обеспечила устойчивость системы к вариациям качества изображений. Выбор MobileNet V3 был обусловлен ее вычислительной эффективностью, что делает модель пригодной не только для лабораторных условий, но и для потенциального применения на устройствах с ограниченными ресурсами, таких как портативные диагностические системы.

Процесс проектирования и обучения модели включал несколько ключевых этапов: анализ предметной области, обзор существующих решений, изучение структуры человеческого глаза и тонометрии, а также разработку архитектуры нейронной сети. Реализация кода на языке Python с использованием фреймворка PyTorch позволила гибко настроить процесс обучения, оптимизировать параметры и сохранить лучшую модель для дальнейшего использования. Применение предварительно обученных весов и методов аугментации данных способствовало достижению приемлемой точности классификации, а использование планировщика скорости обучения и оптимизатора Adam обеспечило стабильность и сходимость модели в течение шести эпох.

Результатом работы стала обученная модель, сохраненная в файл model-eyepac.pth, которая демонстрирует способность эффективно различать изображения глазного дна с признаками глаукомы от нормальных. Тестовая точность, полученная в ходе экспериментов, подтверждает перспективность предложенного подхода, хотя конкретные значения зависят от особенностей данных и могут быть улучшены при дальнейшем увеличении объема обучающей выборки или использовании более сложных архитектур. Кроме того, разработанная система обладает потенциалом для интеграции в клиническую практику, что может значительно упростить процесс скрининга глаукомы и повысить его доступность в регионах с недостатком квалифицированных специалистов.

Несмотря на достигнутые результаты, данная работа имеет возможности для дальнейшего совершенствования. Например, добавление числовых данных тонометрии в качестве дополнительных входных параметров могло бы повысить точность диагностики за счет учета не только визуальных, но и физиологических признаков. Также стоит рассмотреть применение ансамблевых методов или более глубоких архитектур, таких как EfficientNet, при наличии достаточных вычислительных ресурсов. Интерпретируемость модели остается важным направлением для доработки, поскольку врачи нуждаются в понимании, какие признаки изображения повлияли на решение сети, чтобы доверять ее выводам в реальной практике.

Таким образом, данная дипломная работа демонстрирует успешное применение глубокого обучения для классификации глаукомы, подчеркивая потенциал автоматизированных систем в медицине. Полученные результаты могут стать основой для создания более сложных диагностических инструментов, способных интегрироваться в существующие медицинские платформы. В перспективе разработка подобных систем может способствовать снижению заболеваемости слепотой, улучшению качества жизни пациентов и оптимизации работы офтальмологических служб. Исследование подчеркивает важность междисциплинарного подхода, объединяющего медицину и технологии, и открывает путь для дальнейших научных и практических изысканий в этой области.

**СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ**

1. Всемирная организация здравоохранения (ВОЗ). "Глаукома: факты и статистика". [Электронный ресурс]. Режим доступа: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment — Дата доступа: 15.03.2025.
2. Kaggle. "Glaucoma Dataset EyePACS AIROGS Light v2". [Электронный ресурс]. Режим доступа: https://www.kaggle.com/datasets/deathtrooper/glaucoma-dataset-eyepacs-airogs-light-v2/data — Дата доступа: 15.03.2025.
3. Torchvision Documentation. "MobileNet V3 Large Model". [Электронный ресурс]. Режим доступа: https://pytorch.org/vision/stable/models.html#mobilenet-v3 — Дата доступа: 15.03.2025.
4. Официальный сайт Python. [Электронный ресурс]. Режим доступа: https://www.python.org/ — Дата доступа: 15.03.2025.