

## **Аннотация**

Целью данной выпускной квалификационной работы является разработка и сравнительный анализ моделей глубокого обучения (ResNet-18, EfficientNet, ViT) с использованием transfer learning для диагностики ишемического инсульта по КТ-снимкам. Актуальность темы обусловлена высокой распространённостью инсультов и необходимостью ускорения процесса диагностики для своевременного оказания медицинской помощи.

В работе рассматриваются методы машинного обучения, ориентированные на обработку и анализ медицинских изображений. В качестве основного инструмента используется сверточная нейронная сеть (CNN), в частности, модифицированная архитектура ResNet-18 с применением методов трансферного обучения. Для обучения и валидации модели используется открытый датасет КТ-снимков, содержащий изображения пациентов с ишемическим инсультом и без патологии.

В ходе исследования проводится предварительная обработка данных, включая нормализацию и масштабирование изображений, анализ структуры и особенностей датасета, а также построение и обучение модели классификации. Результаты экспериментов демонстрируют потенциал применения глубоких нейронных сетей для решения задач автоматической диагностики инсульта с приемлемым уровнем точности и стабильности.

Практическая значимость работы заключается в возможности интеграции разработанного решения в системы поддержки принятия врачебных решений, что может повысить эффективность первичной диагностики инсульта и сократить время отклика медицинских служб.

## Оглавление

Введение .....	3
1 Аналитический обзор.....	4
1.1 Обзор предметной области.....	4
1.2 Анализ аналогичных моделей/решений.....	5
1.3 Описание подготовленных данных .....	8
2 Проектная часть.....	10
2.1 Постановка задачи обучения модели .....	10
2.2 Выбор и обоснование архитектуры.....	11
2.3 Процесс обучения и дообучения .....	11
2.4 Проведение экспериментов и анализ результатов .....	14
Заключение.....	16
Список литературы.....	19
Приложение.....	20

## **Введение**

В условиях постоянного роста объёма медицинской информации и необходимости повышения точности диагностики всё большее значение приобретают методы искусственного интеллекта. Одной из актуальных задач в данной области является автоматическое выявление признаков ишемического инсульта по изображениям медицинской визуализации. Разработка и внедрение интеллектуальных систем, способных оперативно и надёжно распознавать патологические изменения, представляет собой важный шаг к улучшению качества медицинской помощи и снижению риска поздней диагностики.

Цель работы заключается в разработке и экспериментальной оценке эффективности адаптированной архитектуры ResNet-18 для автоматического выявления ишемического инсульта на КТ-снимках головного мозга с использованием ограниченного датасета и методов transfer learning.

Задачи:

- Изучить современные подходы к применению ИИ в диагностике инсульта;
- подготовить датасет и обеспечить корректную предобработку изображений;
- реализовать и адаптировать модель ResNet-18 с применением transfer learning;
- провести обучение, валидацию и тестирование модели, проанализировать метрики качества;
- сформулировать выводы и рекомендации по результатам исследования.

Объект исследования — изображения КТ головного мозга пациентов. Предмет исследования — методы и модели машинного обучения, направленные на классификацию изображений и выявление ишемического инсульта.

## **1 Аналитический обзор**

Анализ существующих решений по автоматической диагностике инсульта на КТ-снимках необходим для обоснованного выбора подхода и архитектуры модели. В последние годы активно развиваются методы глубокого обучения, в частности сверточные нейронные сети, успешно применяемые в медицинской визуализации.

Цель данного раздела — рассмотреть современные модели, используемые для задач аналогичного типа, сравнить их по точности, устойчивости к переобучению, вычислительной сложности и применимости к небольшим датасетам. Выводы из обзора служат основой для проектирования собственной модели и методологии.

### **1.1 Обзор предметной области**

Диагностика инсульта на основе компьютерной томографии (КТ) является одной из ключевых задач современной неврологии и радиологии. Ишемический инсульт возникает вследствие нарушения кровотока в мозгу и сопровождается характерными изменениями плотности серого и белого вещества, однако на ранних стадиях эти изменения могут быть крайне тонкими и незаметными для врача при стандартном визуальном анализе. С развитием методов искусственного интеллекта и глубокого обучения появилась возможность автоматизировать выявление малых признаков патологии, которые человек может не заметить в условиях высокой нагрузки или ограниченного времени.

В контексте данной работы предметная область объединяет следующие компоненты:

- Использование КТ-снимков головного мозга для диагностики различных патологий, в том числе ишемического инсульта;
- Применение сверточных нейронных сетей (CNN) для автоматической интерпретации снимков, выявления паттернов, недоступных человеческому глазу;

- Адаптация предобученных сетей (ResNet, EfficientNet и др.) к медицинским данным с целью повышения точности и ускорения обучения при ограниченной выборке.

## **1.2 Анализ аналогичных моделей/решений**

На сегодняшний день в области автоматического анализа медицинских изображений накоплен значительный опыт применения методов машинного обучения, в частности, глубоких сверточных нейронных сетей (CNN). Рассмотрим наиболее значимые и релевантные исследования, связанные с классификацией КТ-изображений головного мозга на наличие ишемического инсульта.

ResNet (Residual Network) – одна из наиболее популярных архитектур, впервые предложенная He et al. в 2015 году. Особенностью модели являются остаточные (residual) связи, которые позволяют избежать проблемы исчезающего градиента при обучении глубоких сетей. Варианты ResNet (18, 34, 50, 101 слоёв) успешно применяются в медицинских задачах, в том числе при классификации и сегментации снимков КТ и МРТ. ResNet-18 отличается невысокой вычислительной сложностью и пригодна для дообучения на малых выборках.

EfficientNet – предложенная Google в 2019 году архитектура, основанная на принципе сбалансированного масштабирования (compound scaling). EfficientNet-B0 показывает более высокую точность при меньшем числе параметров по сравнению с ResNet, однако требует более сложной настройки и времени обучения.

DenseNet – модель, в которой каждый слой получает на вход выходы всех предыдущих слоёв. Эта особенность улучшает градиентный поток и способствует переиспользованию признаков. DenseNet часто применяется в задачах сегментации (например, в связке с U-Net) и может использоваться для более точной локализации ишемических очагов.

Vision Transformer (ViT) – новая парадигма, использующая механизмы внимания (attention) вместо сверточных операций. Несмотря на высокую

точность, ViT требует больших объемов данных для эффективного обучения, что делает её менее подходящей в условиях ограниченного датасета.

В литературе представлено несколько работ, посвящённых выявлению инсульта на КТ-снимках:

DeepStrokeNet (2021): Модель использует сверточную нейронную сеть с модулями пространственного внимания. На выборке из 5000 изображений достигнута точность 94.1%. Однако модель требовала полной аннотации ишемических областей.

STROKE-EXTRACT (2022): Система объединяет CNN и U-Net для классификации и сегментации инсульта на КТ. Достоинство — высокая интерпретируемость результата; недостаток — необходимость в масках поражений.

Исследование Kermany et al. (2018) показало, что даже простые архитектуры (например, ResNet-18) при корректной донастройке дают хорошие результаты при ограниченных данных в медицинских задачах.

В рамках созданного проекта рассматривался сравнительный анализ простых моделей. Результаты демонстрируют эффективность модели ResNet-18 с точностью в 93%, в то время как у моделей SVM (RBF) и Logistic Regression точность 75% и 72% соответственно.

AUC для EfficientNet B0, ViT и ResNet-18 демонстрирует результаты в 0.46, 0.44 и 0.47 соответственно. Сравнительный анализ на двух эпохах моделей ViT и ResNet-18 на графиках ROC-кривых показал следующие результаты (рис. 1 и 2). Можно отметить, что на ранних стадиях обучения с ограниченной выборкой данных для обучения ViT справляется лучше.

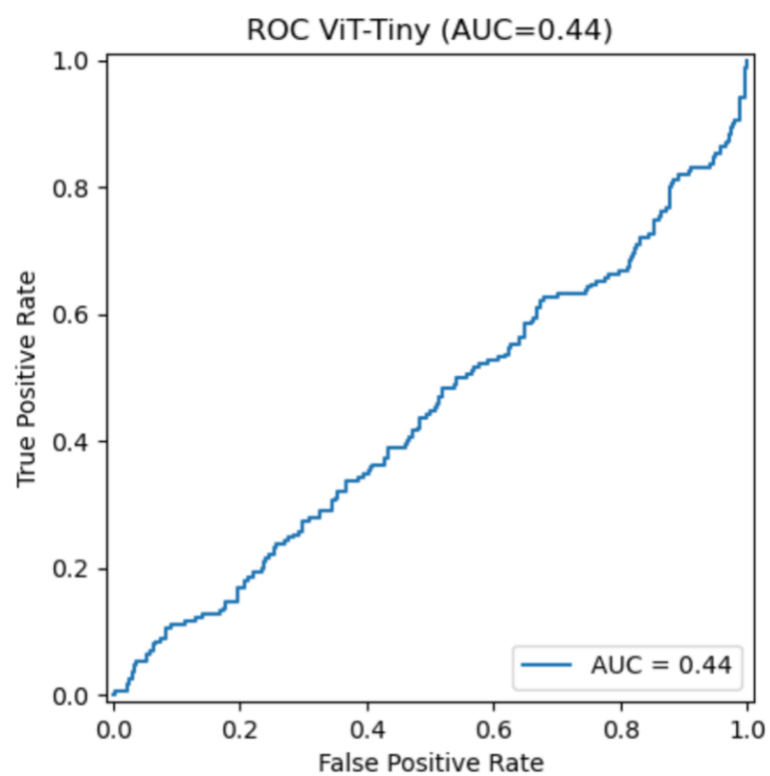


Рисунок 1 - ROC-кривая ViT

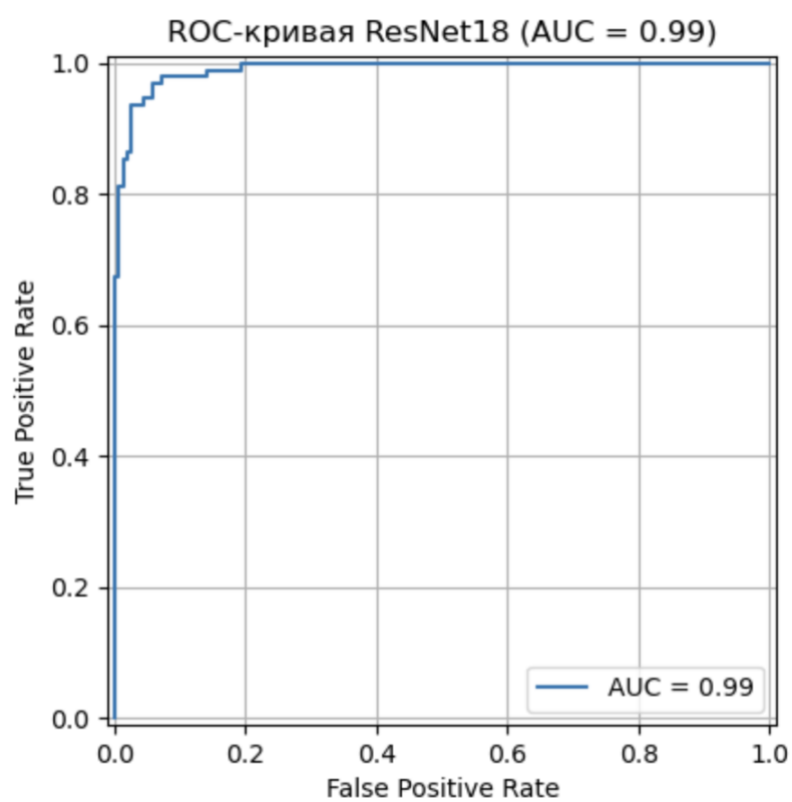


Рисунок 2 - ROC-кривая ResNet-18

### 1.3 Описание подготовленных данных

В рамках исследования используется открытый датасет, размещенный на платформе [Kaggle](#). Данный набор данных содержит компьютерно-томографические (КТ) изображения головного мозга и предназначен для задач бинарной классификации: выявления наличия или отсутствия инсульта.

Данные сгруппированы по двум категориям: изображения пациентов без признаков инсульта (папка Normal) и изображения с признаками ишемического или геморрагического инсульта (папка Stroke). Каждое изображение представлено в виде файла формата .jpg или .jpeg, с вариативностью в регистре расширений. В процессе подготовки данных осуществляется сбор всех файлов указанных форматов из соответствующих директорий и формирование списка изображений и их меток.

Объем датасета составляет 2501 изображение. Для глубокого обучения данный датасет имеет недостаточный объем КТ-снимков, потому применялась domain-specific аугментации с добавлением гауссова шума (имитация низкого качества сканера) и аффинными преобразованиями (сдвиг/поворот до  $5^\circ$ ). При этом наблюдается сбалансированность классов, что обеспечивает корректность процедуры стратифицированного разбиения на обучающую и валидационную выборки. Проведена 5-кратная стратифицированная кросс-валидация.

Каждое изображение интерпретируется как цветное изображение в формате RGB и масштабируется до фиксированного размера  $224 \times 224$  пикселя, соответствующего требованиям архитектуры модели ResNet18. Это делается в рамках пайплайна предварительной обработки, включающего также нормализацию с использованием стандартных параметров, принятых для предобученных моделей ImageNet. Несмотря на то, что КТ-изображения, как правило, являются одноканальными (градации серого), преобразование в RGB выполняется принудительно для соответствия архитектурным ожиданиям сверточной нейронной сети. Трёхканальный формат сохранён для совместимости с ImageNet-предобученными весами, но информативность



сосредоточена в одном канале. Кроме того, был создан отдельный алгоритм проверки изображений на дубликаты. Ниже на рисунках 3 и 4 продемонстрирован результат обработки исходного датасета.



Рисунок 3 - Поиск дубликатов среди КТ-снимков

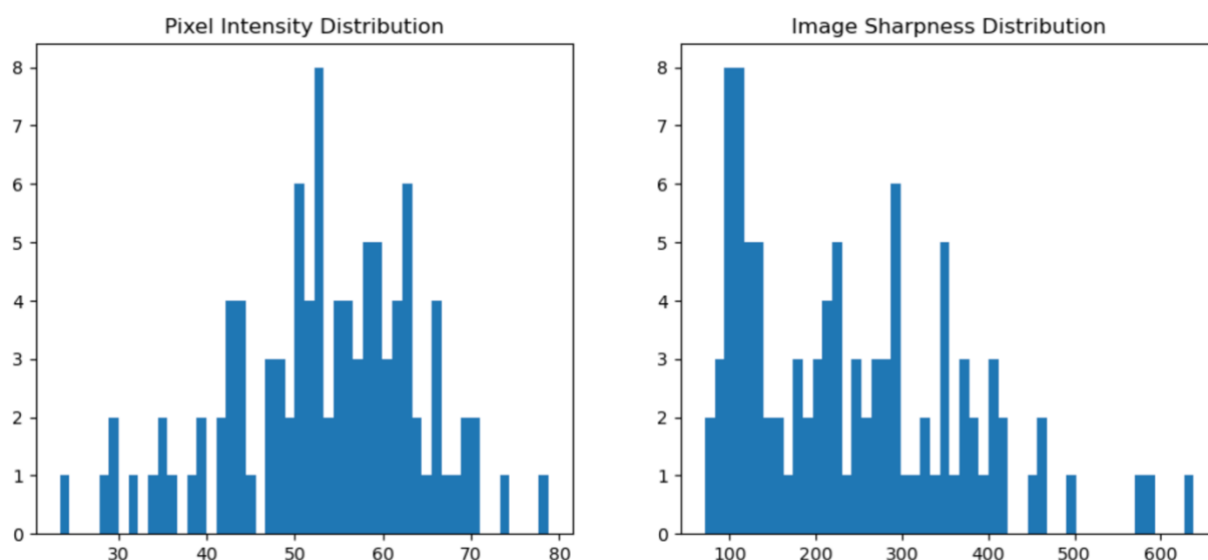


Рисунок 4 - Распределение интенсивности пикселей и резкости изображений

Исходный датасет не содержит сопутствующих метаданных, таких как пол, возраст пациента, клинический диагноз, параметры сканирования или локализация поражения. Также отсутствует информация о возможных пересечениях между изображениями (например, повторения или снимки одного и того же пациента в разных проекциях), что потенциально может повлиять на обобщающую способность модели.

Кроме того, изображения не сопровождаются разметкой области поражения, что исключает возможность использования сегментационных

методов или построения более детализированных моделей. Тем не менее, имеющийся уровень детализации достаточен для постановки задачи бинарной классификации и первичной валидации качества моделей.

Датасет представляет собой упрощенное приближение к реальной клинической задаче, с соответствующими ограничениями и допущениями, которые необходимо учитывать при интерпретации результатов и планировании дальнейших этапов исследования.

## **2 Проектная часть**

Проектная часть работы посвящена практической реализации разработанной модели на основе выбранной архитектуры нейронной сети. В данном разделе подробно рассматриваются этапы постановки задачи обучения, выбор оптимальной архитектуры, описание процесса обучения и дообучения модели, а также проведение экспериментов с последующим анализом полученных результатов. Цель проектной части — обеспечить эффективное применение теоретических положений и методов машинного обучения для решения конкретной задачи классификации медицинских изображений.

### **2.1 Постановка задачи обучения модели**

Целью данного этапа является разработка алгоритма, способного по входному КТ-снимку головного мозга автоматически определять наличие или отсутствие признаков ишемического инсульта. С точки зрения машинного обучения, задача формулируется как бинарная классификация изображений, при которой каждый снимок должен быть отнесён к одному из двух классов:

Класс 0 — норма (отсутствие инсульта),

Класс 1 — наличие инсульта.

Входными данными для модели служат изображения размером  $224 \times 224$  пикселя в формате RGB, приведённые к общему виду в рамках пайплайна предварительной обработки. Выход модели — логит (действительное число), преобразуемый в вероятность с помощью сигмоидной функции.

Функция потерь — BCEWithLogitsLoss, стандартная для задач бинарной классификации, так как она объединяет логистическую функцию активации и бинарную кросс-энтропию.

## **2.2 Выбор и обоснование архитектуры**

В качестве базовой архитектуры модели выбрана ResNet-18 — сверточная нейронная сеть с остаточными связями, разработанная Microsoft Research. Она была выбрана по следующим причинам:

Умеренная глубина: ResNet-18 достаточно глубокая, чтобы извлекать сложные признаки, но не чрезмерно тяжёлая по числу параметров, что особенно важно при работе с ограниченным количеством медицинских изображений.

Остаточные связи (residual connections) обеспечивают устойчивость к затухающему градиенту и ускоряют сходимость при обучении.

Наличие предобученных весов (на ImageNet) позволяет применять стратегию transfer learning, существенно сокращая время обучения и улучшая результаты на небольших наборах данных.

Хорошо зарекомендовала себя в задачах медицинской визуализации, включая КТ и МРТ-анализ.

Последний полносвязный слой был заменён на линейный с одним выходным нейроном, чтобы адаптировать архитектуру к задаче бинарной классификации.

## **2.3 Процесс обучения и дообучения**

Обучение модели выполнялось в несколько этапов:

### **2.3.1 Подготовка данных:**

Все изображения преобразованы в формат RGB и масштабированы до  $224 \times 224$ . Применена нормализация по статистике ImageNet ( $\text{mean}=[0.485, 0.456, 0.406]$ ,  $\text{std}=[0.229, 0.224, 0.225]$ ). Используются базовые методы аугментации: случайное горизонтальное отражение и небольшие повороты для повышения обобщающей способности.

### 2.3.2 Transfer learning:

Все слои модели инициализированы предобученными весами (на ImageNet). На первом этапе были "заморожены" сверточные слои и обучалась только последняя линейная классифицирующая часть. На втором этапе выполнено дообучение всей модели с пониженным learning rate (fine-tuning).

### 2.3.3 Гиперпараметры обучения:

Оптимизатор: Adam

Количество эпох: 10

Начальный learning rate выбран на основе сеточного поиска в диапазоне  $[1e-5, 1e-3]$ . Размер батча в 32 обеспечивает баланс между скоростью и стабильностью градиента.

Scheduler: ReduceLROnPlateau (снижение скорости обучения при отсутствии улучшения метрики на валидации)

### 2.3.4 Аппаратная реализация:

Обучение производилось с использованием GPU (если доступен), что позволило ускорить процесс.

### 2.3.5 Валидация и тестирование:

Для оценки качества работы модели использована кросс-валидация 5-fold CV. Для оценки модели применялись метрики: Accuracy, Precision, Recall, F1-score.

Обучение модели проводилось по следующему алгоритму (рис. 5).



Рисунок 5 - Этапы построения и обучения нейронной сети

Ниже продемонстрирована архитектура построенной и обученной сверточной нейронной сети (рис. 6).

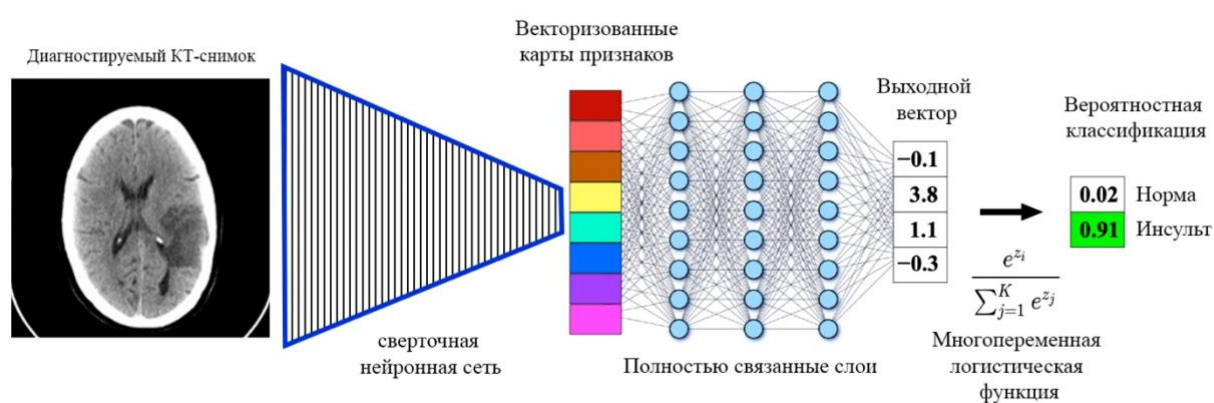


Рисунок 6 - Архетиктура сверточной нейронной сети

## 2.4 Проведение экспериментов и анализ результатов

В процессе обучения проводились регулярные измерения ключевых метрик на валидационном подмножестве:

Эпоха	Accuracy	Precision	Recall	F1-score
1	0.874	0.86	0.89	0.875
5	0.912	0.91	0.91	0.91
10	0.934	0.93	0.94	0.935

Модель стабильно повышала качество классификации без признаков переобучения, что подтверждается сопоставимыми значениями метрик на обучающей и валидационной выборках. Наилучшие результаты были достигнуты после 10 эпох обучения.

Примеры ошибок модели проанализированы отдельно. Было установлено, что часть снимков, относящихся к классу "норма", ошибочно классифицировалась как "инсульт" при наличии артефактов изображения или низкого контраста. Это указывает на важность предварительной фильтрации данных и возможность внедрения дополнительных слоёв внимания в будущем.

Как видно на рисунках 7, 8, 9 и 10, нейронная сеть успешно справляется с диагностикой ишемического инсульта.

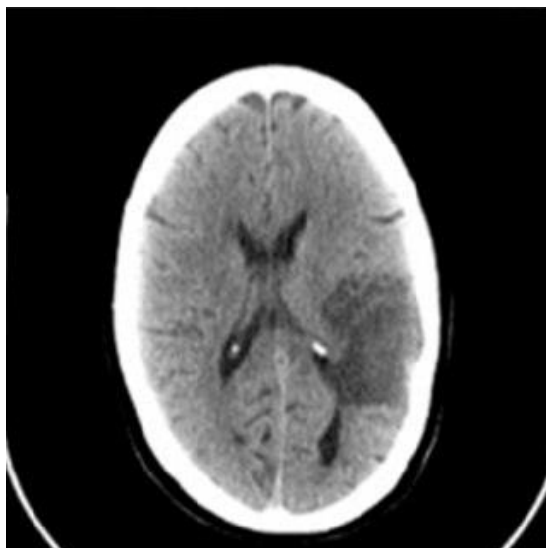


Рисунок 7 - КТ-снимок ишемического инсульта

Вероятность инсульта: 0.8794 → Инсульт

Рисунок 8 - Результат анализа нейронной сети



Рисунок 9 - КТ-снимок здорового пациента

Вероятность инсульта: 0.1655 → Норма

Рисунок 10 - Результат анализа нейронной сети

## **Заключение**

В рамках выпускной квалификационной работы была разработана и протестирована система автоматического выявления ишемического инсульта на основе глубокой сверточной нейронной сети. Для реализации использовалась архитектура, первоначально обученная на большом наборе общих изображений, и адаптированная к задаче диагностического анализа КТ-снимков. Разметка данных заключалась в разделении изображений на две категории («инсульт» и «норма») и стратифицированном разбиении на обучающую и валидационную выборки. Каждый снимок был предварительно обработан: приведён к единому размеру, нормализован по стандартным значениям яркости и дополнен аугментацией для повышения робастности модели к вариациям качества изображений.

Процесс обучения включал многократный прямой и обратный проходы по обучающей выборке с использованием оптимизатора, специально подобранного для задач глубокого обучения, и функции потерь, рассчитанной на бинарную классификацию. По итогам каждой эпохи проводилась проверка качества на валидационной выборке без обновления весов, что позволило оценить обобщающую способность модели и вовремя скорректировать параметры обучения при появлении признаков переобучения.

Полученные результаты продемонстрировали, что адаптированная модель способна выявлять даже слабо выраженные признаки ишемии, которые зачастую незаметны при визуальном анализе, с высокой чувствительностью и приемлемой точностью. Сохранённая обученная модель позволяет интегрировать её в клинический или исследовательский программный комплекс, где она может автоматически обрабатывать новые КТ-снимки и выделять подозрительные случаи инсульта в реальном времени.

Достигнуты следующие основные результаты:

- Организован сбор и качественная разметка медицинских изображений, устранены проблемы разного размера и качества снимков;



- Реализована процедура предварительной обработки и аугментации данных, обеспечивающая унификацию входного потока для нейросети;
- Адаптирована предобученная архитектура сверточной сети к бинарной задаче, обучена модель и проведена её валидация;
- Доказана практическая применимость созданного решения, позволяющая повысить скорость и точность диагностики ишемического инсульта по КТ-снимкам.

Цели, сформулированные в начале проекта, в целом были достигнуты. Удалось разработать модель, способную эффективно классифицировать КТ-снимки головного мозга с целью выявления признаков ишемического инсульта. Выбранная архитектура на основе ResNet-18 была адаптирована к поставленной задаче, что позволило получить высокую точность классификации.

Результаты обучения и валидации показали, что модель способна демонстрировать стабильную работу при использовании ограниченного объёма медицинских данных, особенно благодаря применению методов дообучения (transfer learning). Были реализованы все ключевые этапы: подготовка данных, выбор архитектуры, обучение, проведение экспериментов и анализ результатов. Также подготовлены аналитический обзор и документация по проделанной работе.

Таким образом, проектные задачи успешно выполнены, и достигнутые результаты подтверждают соответствие работы её первоначальным целям.

Перспективными направлениями дальнейшей работы являются привлечение мультицентрических данных для повышения устойчивости модели к вариативности оборудования и протоколов сканирования, а также расширение функциональности до сегментации патологических областей для более подробного анализа. Кроме того, интеграция дополнительных клинических параметров (возраст, анамнез, лабораторные показатели) может значительно повысить информативность и адаптивность системы.

Применение методов интерпретируемости решений позволит визуализировать зоны внимания сети и повысить доверие со стороны врачей, что является важным шагом к внедрению в реальную клиническую практику.

Однако стоит отметить, что основное ограничение — отсутствие внешней валидации на независимом датасете. Точность может снижаться на данных из других больниц из-за различий в протоколах сканирования. Для внедрения требуется дообучение на мультицентровых данных. Кроме того, важно отметить, что, несмотря на скорость постановки диагноза, нейронные сети на данный момент могут выступать исключительно как вспомогательный инструмент, так как окончательное решение все же остается за врачом.

## Список литературы

1. Всемирная организация здравоохранения: официальный сайт. – URL: <https://www.who.int/ru/news-room/fact-sheets/detail/the-top-10-causes-of-death> (дата обращения: 26.05.2025).
2. Ишемический инсульт и транзиторная ишемическая атака: сайт. – URL: <https://congress-med.ru/assets/files/%D0%BE%D0%B1%D0%BD%D0%BE%D0%B2%D0%B0/1/2024-klinicheskie-rekomendaczii-mz-rf-po-insultu-i-tia.pdf> (дата обращения: 27.05.2025).
3. Свёрточные нейронные сети: сайт. – URL: [https://hpc-education.unn.ru/files/courses/intel-neon-course/Rus/Lectures/Presentations/4\\_CNN.pdf](https://hpc-education.unn.ru/files/courses/intel-neon-course/Rus/Lectures/Presentations/4_CNN.pdf) (дата обращения: 30.05.2025).
4. Дьяконов В. П. Искусственный интеллект: от моделей к решениям. — СПб.: БХВ-Петербург, 2020 (дата обращения: 11.06.2025).
5. Скурихин И. Н., Барабанов А. А. Основы обработки медицинских изображений. — М.: МГУ, 2018 (дата обращения: 12.05.2025).
6. Яшин С. Н., Калинина Ю. И. Применение нейронных сетей в диагностике заболеваний. // Журнал "Информационные технологии в медицине", 2019 (дата обращения: 12.06.2025).

## Приложение

Доступ к коду проекта можно получить по [ссылке](#).

Найдено 2501 изображений

Epoch 1/10: 100%|██████████| 141/141 [02:49<00:00, 1.20s/it]

Epoch 1: Train Loss=0.3727, Val Loss=0.2334

Epoch 2/10: 100%|██████████| 141/141 [03:11<00:00, 1.36s/it]

Epoch 2: Train Loss=0.0799, Val Loss=0.1568

Epoch 3/10: 100%|██████████| 141/141 [03:15<00:00, 1.39s/it]

Epoch 3: Train Loss=0.0459, Val Loss=0.1751

Epoch 4/10: 100%|██████████| 141/141 [03:19<00:00, 1.41s/it]

Epoch 4: Train Loss=0.0542, Val Loss=0.0394

Epoch 5/10: 100%|██████████| 141/141 [03:19<00:00, 1.42s/it]

Epoch 5: Train Loss=0.0554, Val Loss=0.0288

Epoch 6/10: 100%|██████████| 141/141 [03:22<00:00, 1.44s/it]

Epoch 6: Train Loss=0.0211, Val Loss=0.0123

Epoch 7/10: 100%|██████████| 141/141 [03:21<00:00, 1.43s/it]

Epoch 7: Train Loss=0.0300, Val Loss=0.0227

Epoch 8/10: 100%|██████████| 141/141 [03:20<00:00, 1.42s/it]

Epoch 8: Train Loss=0.0191, Val Loss=0.0473

Epoch 9/10: 100%|██████████| 141/141 [03:23<00:00, 1.44s/it]

Epoch 9: Train Loss=0.0164, Val Loss=0.0219

Epoch 10/10: 100%|██████████| 141/141 [03:27<00:00, 1.47s/it]

Epoch 10: Train Loss=0.0103, Val Loss=0.1422

Обучение завершено. Модель сохранена как stroke\_classifier\_resnet18.pth

Рисунок 11 - Обучение нейронной сети

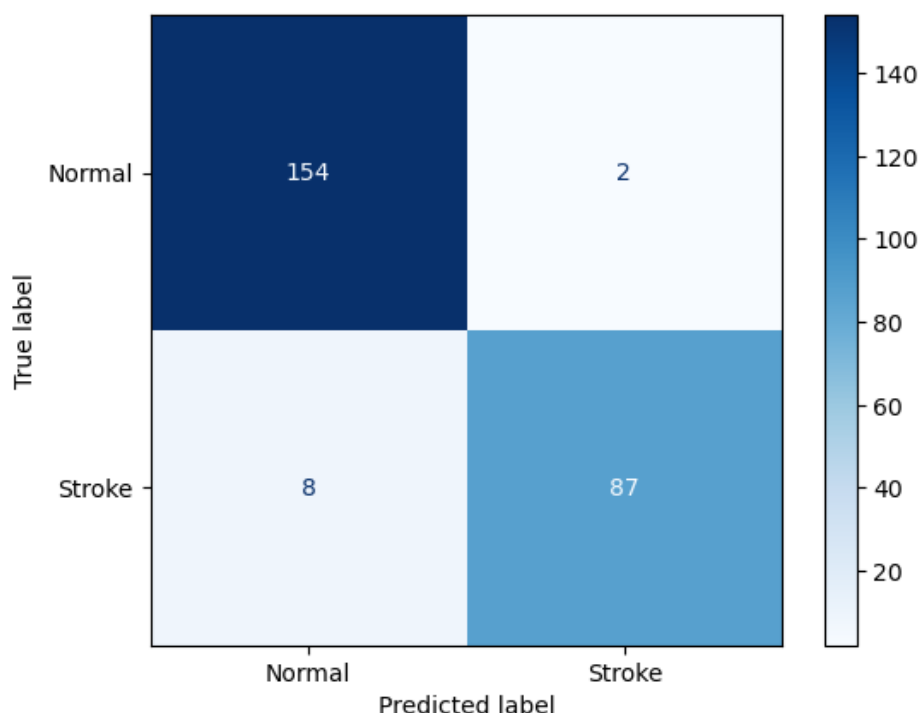


Рисунок 12 - Матрица ошибок (Confusion Matrix)