Министерство образования Республики Беларусь

Учреждение образования «Брестский государственный Технический университет»

Кафедра ИИТ

Пояснительная записка

к курсовой работе по дисциплине

«Модели решения задач в интеллектуальных системах»

по теме

«Разработка и сравнительный анализ моделей машинного обучения для перевода языка жестов ASL в текст»

КР. ИИ-23.220098 40-03-01

Листов: 19

**Выполнил**

студент 3 курса,

ФЭИС, группы ИИ-23

Макаревич Н.Р.

**Нормоконтроль**

Туз И. С.

**Проверил**

Туз И. С.

Брест 2025

Содержание

[ВВЕДЕНИЕ 3](#_Toc197974036)

[Глава 1. ML-Задача и датасет 4](#_Toc197974037)

[1.1.1 Постановка задачи 4](#_Toc197974038)

[1.1.2 Обоснование выбора задачи 4](#_Toc197974039)

[1.1.3 Выбор датасета 5](#_Toc197974040)

[Глава 2 Описание задачи: цели, особенности, метрики, модели 6](#_Toc197974041)

[2.1.1 Цель задачи 6](#_Toc197974042)

[2.1.2 Особенности задачи 6](#_Toc197974043)

[2.1.3 Метрики оценки качества 6](#_Toc197974047)

[2.1.4 Модели, подходящие для решения задачи 6](#_Toc197974048)

[Глава 3. Обзор моделей 7](#_Toc197974049)

[3.1.1 SimpleNN 7](#_Toc197974051)

[3.1.2 SimpleCNN 8](#_Toc197974052)

[3.1.3 AdvancedCNN 8](#_Toc197974053)

[3.1.4 MediapipeNN 12](#_Toc197974054)

[3.2.1 Выводы 14](#_Toc197974083)

[ЗАКЛЮЧЕНИЕ 16](#_Toc197974084)

[СПИСОК ЛИТЕРАТУРЫ 17](#_Toc197974085)

ПРИЛОЖЕНИЕ A. ТЕКСТ программы

# ВВЕДЕНИЕ

Современное развитие искусственного интеллекта (ИИ) и глубинного обучения привело к значительным достижениям в области обработки видеоданных и распознавания человеческих действий. Одной из перспективных и активно исследуемых задач является автоматический перевод языка жестов — в частности, американского языка жестов (American Sign Language, ASL) — в текстовую форму.

Распознавание ASL — это нетривиальная задача, так как включает в себя анализ как статических, так и динамических особенностей движений рук, мимики и телесной позы. Модель должна не только точно распознавать отдельные жесты, но и учитывать их контекст в пределах всей фразы, что требует сложной обработки временной информации и пространственно-временных зависимостей.

В последние годы большое внимание уделяется применению различных архитектур нейросетей к этой задаче, включая сверточные сети (CNN), рекуррентные сети (RNN), трансформеры. Эти подходы позволяют по-разному обрабатывать видеопоток, позные данные (например, координаты суставов) и другие признаки, извлекаемые из видео.

Перевод языка жестов можно рассматривать как задачу преобразования данных из одного высокоразмерного пространства в последовательность символов или слов — аналогично переводу с одного языка на другой. Это требует построения моделей, которые умеют эффективно кодировать визуальные данные и декодировать их в текстовую форму.

Ключевыми аспектами при построении таких моделей являются:

* выбор входного представления (видео, скелетные координаты, смешанные признаки);
* архитектура модели, учитывающая как пространственные, так и временные зависимости;
* выбор обучающей стратегии и потерь, подходящих для языковой модели;
* метрики качества, позволяющие объективно сравнивать модели по точности перевода.

В рамках настоящего курсового проекта ставится задача разработки и сравнительного анализа нескольких подходов к обучению моделей для перевода ASL в текст. Будут реализованы и протестированы пять различных архитектур, охватывающих как базовые, так и современные модели, включая CNN, MLP, и их модификации. Все модели будут обучены и протестированы на специализированных датасетах ASL, а также проанализированы с точки зрения точности, устойчивости и практической применимости.

#### Глава 1. ML-Задача и датасет

## Постановка задачи

Перевод американского языка жестов (ASL) в текст является важной задачей в области компьютерного зрения и обработки естественного языка. Она представляет собой разновидность задачи мультимодального машинного перевода, в которой входными данными являются либо видеопоследовательности с жестами, либо извлечённые из них ключевые признаки (например, координаты суставов рук), а выходными — текстовые предложения на английском языке.

Формально задача может быть описана следующим образом:  
Пусть дана выборка D={(V1,T1),...,(Vn,Tn)}D = \{(V\_1, T\_1), ..., (V\_n, T\_n)\}D={(V1​,T1​),...,(Vn​,Tn​)}, где ViV\_iVi​ — входная видеопоследовательность (или последовательность признаков), содержащая ASL-жесты, а TiT\_iTi​ — соответствующий текст на английском языке. Необходимо построить модель f(V;θ)f(V; \theta)f(V;θ), параметризованную θ\thetaθ, такую что f(Vi)≈Tif(V\_i) \approx T\_if(Vi​)≈Ti​.

В рамках данной работы рассматриваются модели, способные решать задачу сопоставления входного визуального или позового сигнала с соответствующим текстовым представлением. Основной акцент сделан на сравнительном анализе различных архитектур нейронных сетей для повышения точности и устойчивости перевода.

## Обоснование выбора задачи

Задача автоматического перевода ASL в текст обладает высокой практической значимостью и исследовательской ценностью:

* **Социальная значимость:** автоматический перевод жестов помогает преодолеть коммуникативный барьер между людьми с нарушениями слуха и остальным обществом;
* **Сложность задачи:** требует моделирования как пространственной (формы жестов), так и временной (последовательности движений) информации;
* **Отсутствие явных классов:** перевод ASL не сводится к классификации, что делает задачу более гибкой и открытой для архитектур, работающих с последовательностями;
* **Разнообразие подходов:** предоставляет возможность сравнить методы, основанные на CNN, RNN, Transformer, 3D-CNN и других подходах.

## Выбор датасета

В данной работе используется ASL(American Sign Language) Alphabet Dataset, который является открытым и часто используемым в исследованиях по генерации мультипликационных изображений.



Рисунок 1.1 – Пример изображений из датасета

Основные характеристики датасета:

* 223000 изображений;
* Размер изображений: 640 × 480 пикселя, формат JPG;
* Каждое изображение представляет собой жест руки;

Выбор данного датасета обусловлен следующими причинами:

* Большой объем данных;
* Высокое качество и однородность изображений;
* Распространенность исследований, что позволяет сравнивать результаты с другими работами.

# Глава 2 Описание задачи: цели, особенности, метрики, модели

## Цель задачи

Целью рассматриваемой задачи является разработка модели, способной эффективно интерпретировать и переводить жесты языка жестов в текст.

## Особенности задачи

Модель должна удовлетворять следующим критериям:

* Точность распознавания жестов: корректное определение формы рук, ориентации и движения;
* Универсальность: способность работать с разными людьми и условиями съёмки (фон, освещение);
* Низкая задержка: возможность применения в режиме реального времени.

## Метрики оценки качества генерации

Так как задача заключается в распознавании отдельных **букв**, она сводится к **классификации жестов по фиксированному числу классов**. В этом контексте применимы стандартные метрики из области классификации:

* **Accuracy (Точность)**: доля правильно распознанных букв от общего числа. Базовая и наиболее используемая метрика.
* **Precision / Recall / F1-Score**: особенно полезны при наличии дисбаланса между классами (например, если некоторые буквы встречаются чаще других).

## Модели, подходящие для решения задачи

Для задачи распознавания **отдельных букв** из языка жестов целесообразно использовать следующие архитектуры:

* **CNN** (сверточные нейросети): хорошо подходят для распознавания статичных изображений.
* **Keypoint-based модели** (с использованием MediaPipe или OpenPose): сначала извлекаются ключевые точки рук, затем они классифицируются при помощи MLP, LSTM или Transformer.

# Глава 3. Обзор моделей GAN



## SimpleNN

SimpleNN — базовая архитектура многослойного перцептрона (MLP), реализованная с использованием фреймворка PyTorch. Она предназначена для решения задачи классификации изображений, в частности — для распознавания **отдельных букв дактилологии** (пальцевой азбуки) по RGB-кадрам фиксированного размера (64×64 пикселя). Модель отличается простотой конструкции и минимальным числом параметров, что делает её подходящей в качестве базового решения и отправной точки для последующего усложнения.

Модель включает в себя следующие компоненты:

* **Слой Flatten** — преобразует входной тензор формы (3, 64, 64) в одномерный вектор длины 12288 (то есть 3×64×64). Это необходимо для подачи данных на полносвязный слой.
* **Полносвязный слой Linear(12288, 128)** — линейное преобразование входного вектора в пространство признаков размерности 128.
* **Функция активации ReLU** — вводит нелинейность, позволяя модели аппроксимировать сложные зависимости в данных.
* **Выходной слой Linear(128, num\_classes)** — преобразует скрытое представление в логиты размерности num\_classes, соответствующие числу классов (в случае дактилологии — 29: 26 латинских букв + 3 спецсимвола, таких как «delete», «space» и «nothing»).

Функция прямого распространения (forward) последовательно применяет указанные преобразования к входному изображению и возвращает вектор логитов, интерпретируемый как вероятностное распределение по классам (после применения Softmax вне модели).

#### Преимущества модели

* **Простота и прозрачность архитектуры**: модель легко интерпретируема, быстро обучается и подходит для отладки пайплайна.
* **Низкие вычислительные требования**: отсутствие свёрточных слоёв и небольшое число параметров делает её пригодной для обучения даже на CPU.
* **Быстрая итерация**: за счёт малой глубины и размеров модели возможно быстрое проведение экспериментов и подбор гиперпараметров.

#### Ограничения модели

* **Отсутствие пространственной инвариантности**: в отличие от свёрточных сетей, MLP не способен учитывать локальные паттерны и смещения в изображении, что особенно важно при работе с визуальными данными.
* **Склонность к переобучению на малых выборках**: из-за большого входного пространства (более 12 тысяч признаков) и малого числа скрытых узлов модель может либо недообучаться, либо быстро переобучаться, особенно без регуляризации.
* **Плохая масштабируемость**: с увеличением разрешения изображений или количества классов модель потребует непропорционально больше параметров, что ограничивает её применимость для более сложных задач.

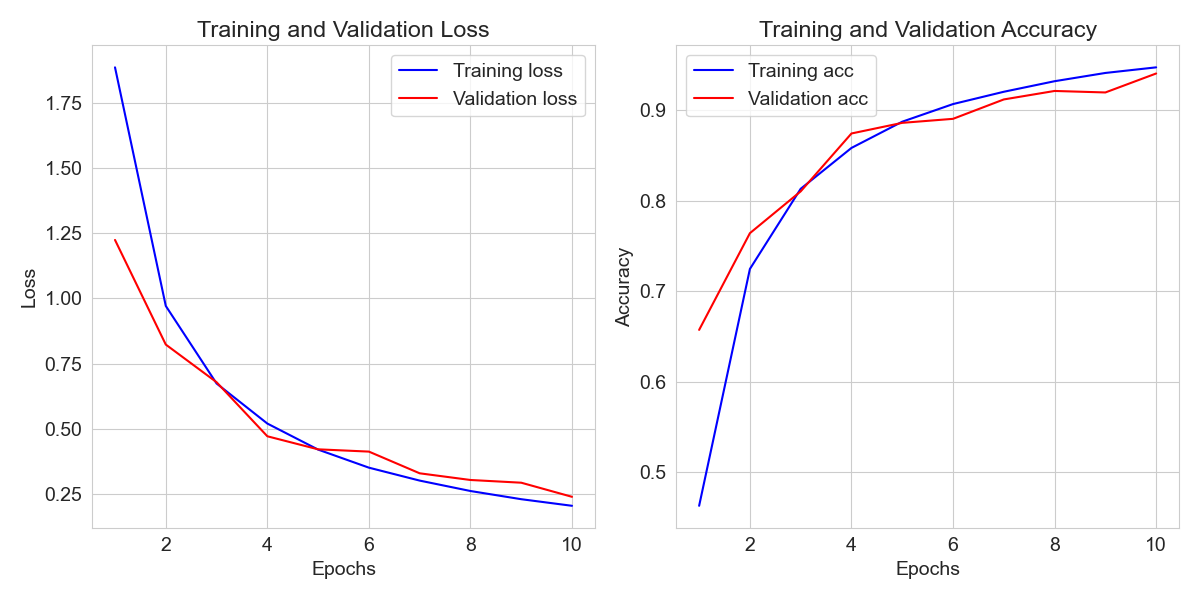


Рисунок 2.1 – результаты обучения сети

## SimpleCNN

SimpleCNN — базовая архитектура сверточной нейронной сети (CNN), предназначенная для классификации изображений жестов, представляющих **отдельные буквы дактилологии** (языка жестов). В отличие от многослойного перцептрона, эта модель учитывает **локальные пространственные зависимости** в изображениях, что значительно повышает её способность извлекать устойчивые к сдвигам и масштабированию признаки.

Архитектура сети состоит из двух основных блоков: **экстрактора признаков** (модуля features) и **классификатора** (classifier):

#### 1. ****Feature extractor****:

* Conv2d(3, 16, 3, padding=1) — первый сверточный слой с 16 фильтрами 3×3, который принимает на вход цветные изображения с тремя каналами (RGB). Padding сохраняет исходный размер пространственного разрешения (64×64).
* ReLU() — функция активации, вводящая нелинейность.
* MaxPool2d(2) — операция субдискретизации, уменьшающая пространственное разрешение в 2 раза (до 32×32).
* Conv2d(16, 32, 3, padding=1) — второй сверточный слой с увеличением числа каналов до 32.
* ReLU() и MaxPool2d(2) — те же операции, в результате которых размер изображения становится 16×16.

#### 2. ****Классификатор****:

* Flatten() — преобразует выход сверточного блока размерности (32, 16, 16) в одномерный вектор длины 8192.
* Linear(8192, 128) — полносвязный слой, сжимающий пространство признаков до размерности 128.
* ReLU() — активация.
* Linear(128, num\_classes) — выходной слой, выдающий логиты по числу классов (в данной задаче — 29 букв).

#### Преимущества модели

* **Учет локальных признаков**: свёрточные фильтры позволяют эффективно извлекать текстуры, границы и формы из изображений рук.
* **Меньшее количество параметров по сравнению с MLP**: благодаря повторному использованию фильтров, CNN легче масштабировать при сохранении вычислительной эффективности.
* **Лучшая обобщающая способность**: за счёт пространственной инвариантности модель устойчивее к вариациям в изображениях, включая изменение масштаба, поворота и освещения.

#### Ограничения модели

* **Ограниченная глубина**: двухуровневая архитектура может оказаться недостаточной для захвата сложных паттернов, особенно при высоком разнообразии жестов и фонов.
* **Фиксированный входной размер**: как и у большинства CNN, размер входа должен быть постоянным (в данном случае — 64×64), что требует предварительной нормализации данных.
* **Зависимость от гиперпараметров**: выбор количества фильтров, размеров ядер и числа слоёв сильно влияет на производительность, и требует тонкой настройки.

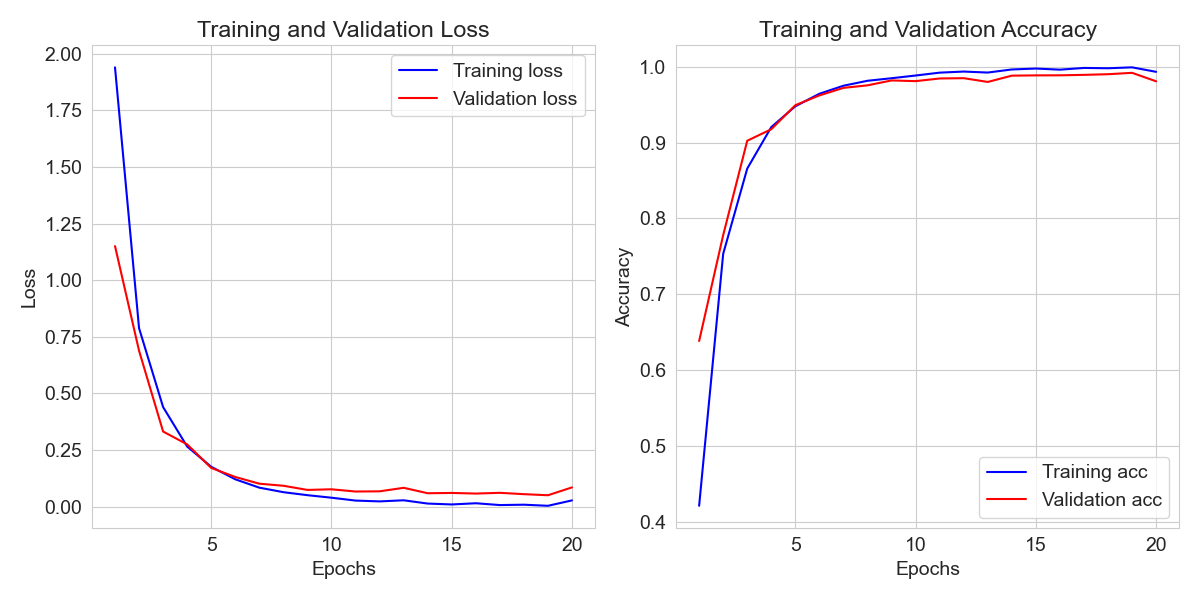


Рисунок 2.2 – результаты обучения сети

## ****AdvancedCNN****

**AdvancedCNN** — усовершенствованная архитектура сверточной нейронной сети, предназначенная для классификации изображений с буквами языка жестов. По сравнению с базовой моделью SimpleCNN, данная сеть отличается **увеличенной глубиной**, **нормализацией активаций** и использованием **адаптивного усреднения**, что позволяет добиться **более стабильного обучения и лучшей обобщающей способности**.

Модель состоит из двух компонентов: **экстрактора признаков** и **классификатора**.

#### 1. ****Feature extractor**** (features)

Состоит из трёх сверточных блоков:

* **Первый блок**:
  + Conv2d(3, 32, 3, padding=1) — извлекает 32 признака из входного изображения.
  + BatchNorm2d(32) — нормализует выход свёртки, ускоряя обучение и повышая устойчивость к переобучению.
  + ReLU() — активация.
  + MaxPool2d(2) — понижение размерности до 32×32.
* **Второй блок**:
  + Conv2d(32, 64, 3, padding=1)
  + BatchNorm2d(64)
  + ReLU()
  + MaxPool2d(2) — размерность становится 16×16.
* **Третий блок**:
  + Conv2d(64, 128, 3, padding=1)
  + BatchNorm2d(128)
  + ReLU()
  + AdaptiveAvgPool2d((1, 1)) — вместо фиксированного MaxPool, используется адаптивное усреднение, которое независимо от входной размерности приводит к выходу 1×1×128.

#### 2. ****Классификатор****

* Linear(128, num\_classes) — принимает вектор признаков длины 128 и предсказывает вероятности по 29 классам (буквам алфавита дактилологии).

#### Преимущества модели

* **Глубокая и стабильная архитектура**: три сверточных слоя с прогрессивным увеличением числа каналов позволяют извлекать иерархические признаки, от простых краёв до абстрактных форм рук.
* **Batch Normalization**: снижает внутреннее ковариационное смещение, ускоряет обучение и повышает устойчивость к переобучению.
* **AdaptiveAvgPool2d**: делает модель независимой от исходного размера входных изображений, что удобно при работе с различными датасетами.
* **Компактный классификатор**: всего один линейный слой снижает риск переобучения и упрощает интерпретацию результатов.

#### Ограничения модели

* **Увеличенное время обучения**: по сравнению с SimpleCNN, из-за большего числа параметров.
* **Отсутствие регуляризации (Dropout)**: при наличии шумных данных может потребоваться добавить методы регуляризации.
* **Всё ещё ограниченная глубина**: по сравнению с современными архитектурами ResNet или EfficientNet, модель остаётся относительно простой.

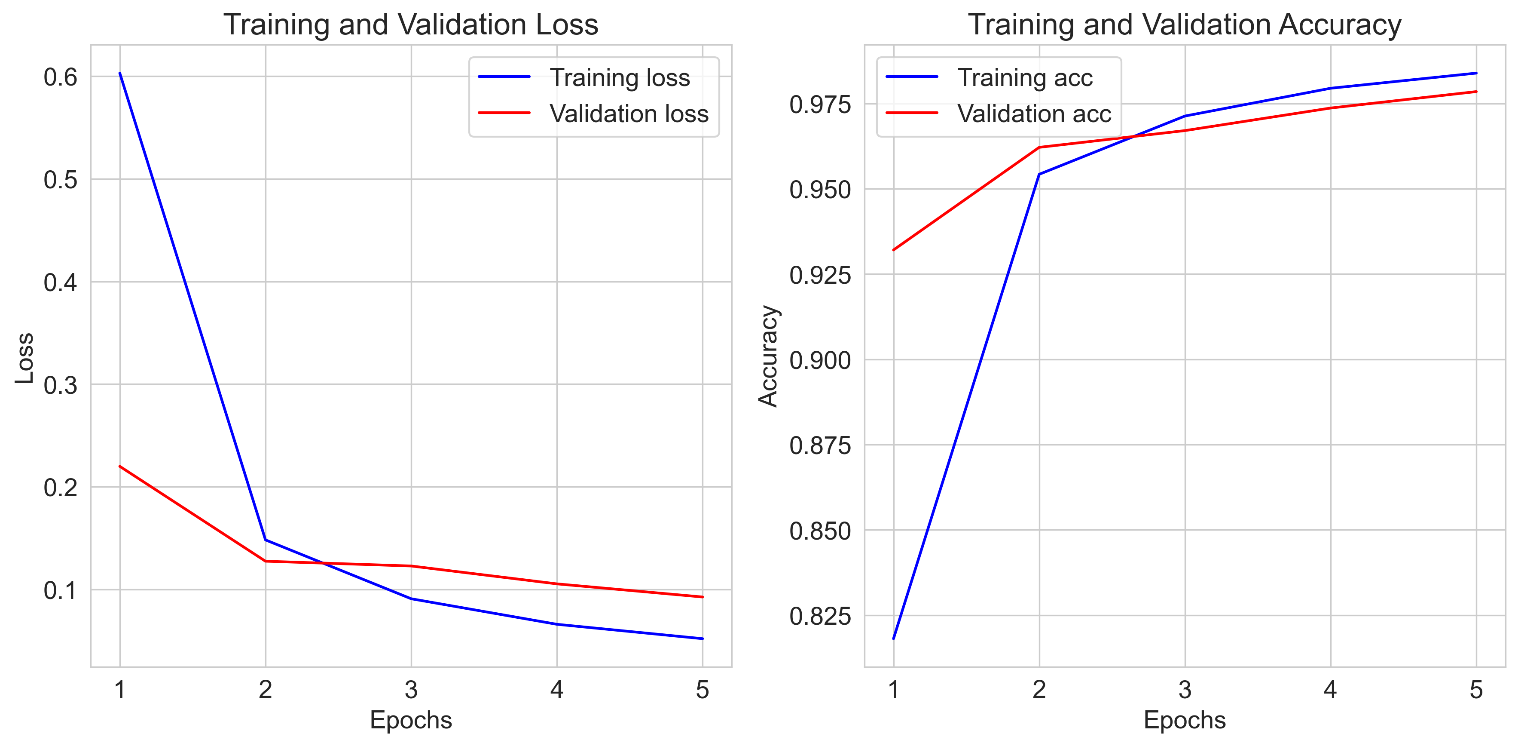


Рисунок 2.3 – результаты обучения сети

## MPNN

**MPNN (MediaPipe Neural Network)** — компактная полносвязная нейронная сеть, специально разработанная для классификации жестов по признакам, извлечённым с помощью фреймворка **MediaPipe**. Вместо работы с изображениями напрямую, модель использует обработанные данные, полученные в результате детекции и анализа положения кисти руки.

#### Предварительная обработка с MediaPipe

Перед подачей на вход модели изображение проходит через **MediaPipe Hands** — инструмент, который:

* Находит кисть руки на изображении,
* Выделяет 21 ключевую точку (landmark) кисти,
* Записывает координаты этих точек (x, y, z) в вектор из **63 чисел** (21 × 3),
* Производит нормализацию координат для унификации данных.

Таким образом, вместо многоканального изображения размером 64×64×3, сеть получает компактный вектор с важной скелетной информацией, существенно сокращая размер входных данных и повышая скорость обучения.

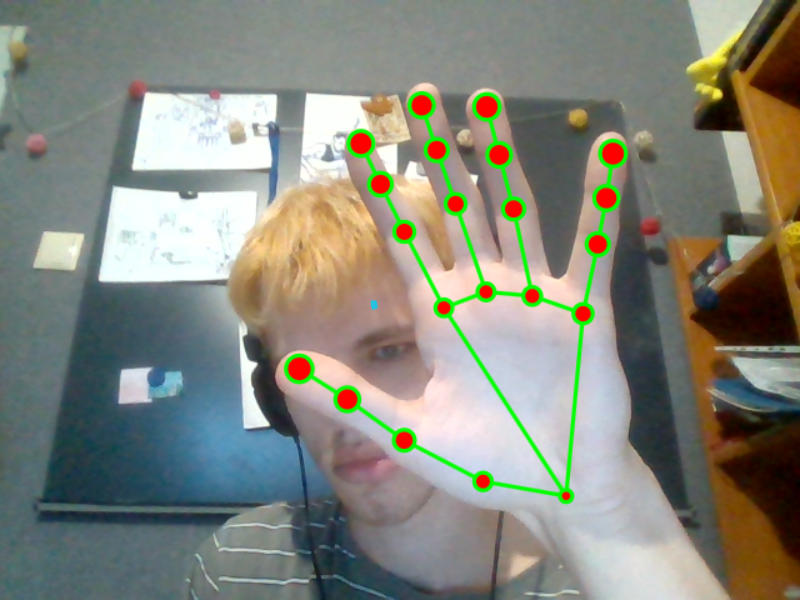


Рисунок 2.4 – результаты предпроцессора

#### Архитектура MPNN

* Входной размер input\_dim равен 63 — числу признаков, получаемых после MediaPipe.
* Сеть состоит из трёх полносвязных слоёв с активацией ReLU, что обеспечивает достаточно мощное, но при этом простое преобразование признаков в вероятности классов.

#### Преимущества MPNN

* **Компактность и скорость:** вектор из 63 признаков гораздо легче и быстрее обрабатывается, чем изображение.
* **Устойчивость к фоновым шумам и вариациям освещения**, так как модель работает не с пикселями, а с координатами ключевых точек.
* **Интерпретируемость:** каждый признак соответствует конкретной точке кисти, что облегчает анализ и отладку модели.

#### Ограничения

* **Зависимость от точности MediaPipe:** ошибки в распознавании и локализации кисти ухудшают итоговые результаты.
* **Потеря визуального контекста:** отсутствуют данные о текстуре, цвете и окружающей среде.
* **Отсутствие временной информации:** анализируется только текущий кадр, без учёта динамики жеста.

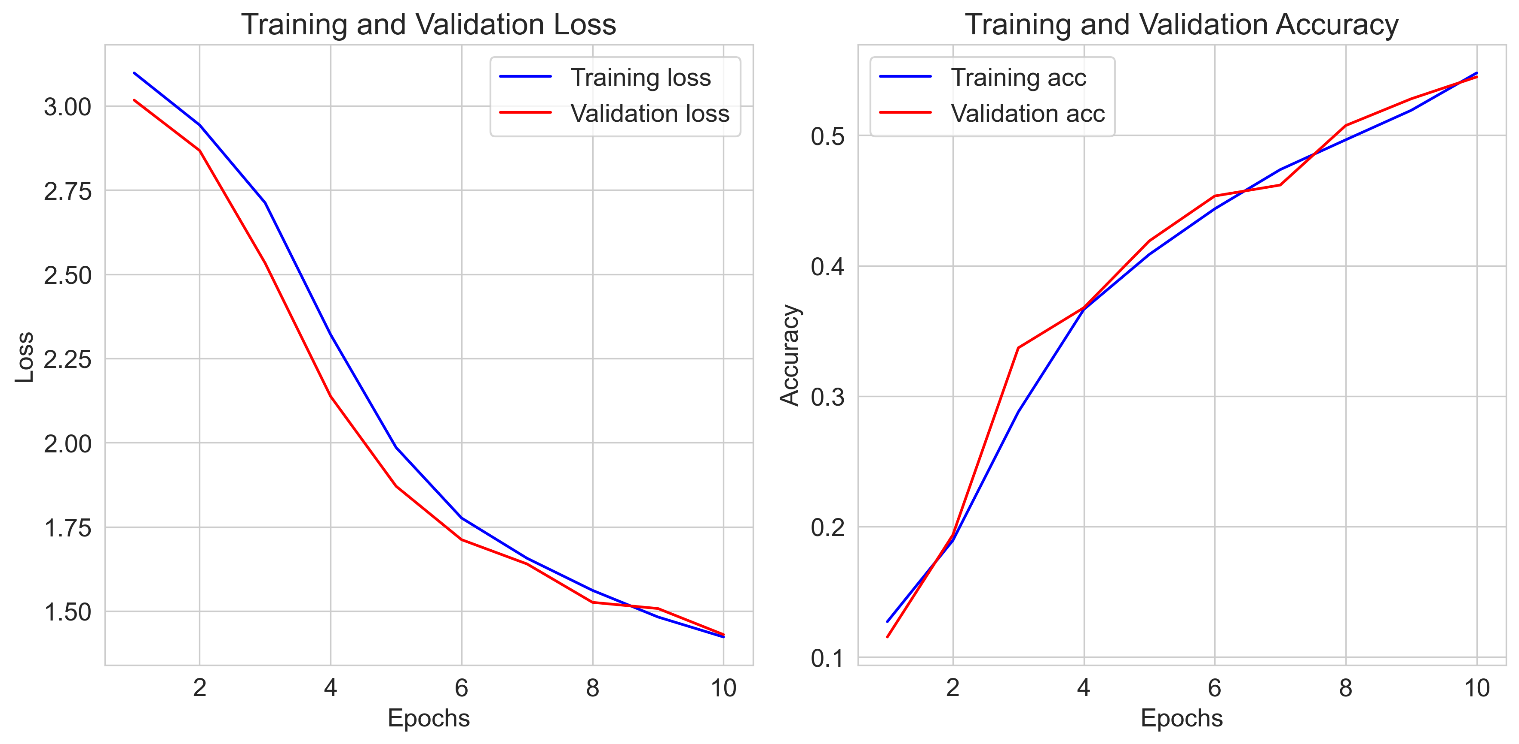


Рисунок 2.5 – результаты обучения сети

## Выводы

SimpleNN — самая базовая архитектура из рассмотренных моделей. Несмотря на простоту и малое количество параметров, она практически не справляется с задачей распознавания букв языка жестов. Отсутствие сверточных слоёв и работы с пространственными признаками приводит к очень низкому качеству классификации. Кроме того, из-за большого входного размера (64×64×3, расплющенного в вектор) и отсутствия эффективной обработки изображений, SimpleNN показывает одни из худших результатов по времени обучения и по точности. Эта модель служит скорее отправной точкой для сравнения и демонстрации важности архитектур с учётом особенностей данных.

SimpleCNN — первая сверточная сеть, которая уже способна эффективно извлекать пространственные признаки из изображений кисти руки. Благодаря двум сверточным слоям с активацией ReLU и подвыборке через MaxPooling, модель показывает существенное улучшение по сравнению с SimpleNN. SimpleCNN достигает приемлемого баланса между скоростью обучения и качеством распознавания, но при этом ограничена по глубине и сложности. Визуальное качество классификации и метрики точности выше, чем у SimpleNN, однако модель всё ещё уступает более продвинутым архитектурам.

AdvancedCNN — более сложная сверточная сеть, использующая дополнительные техники: Batch Normalization, адаптивный pooling и большее число каналов. Эта архитектура демонстрирует заметный прирост точности и устойчивости при обучении, а также улучшенное время сходимости благодаря нормализации. Более глубокие слои и улучшенная обработка признаков позволяют модели распознавать более сложные закономерности жестов, что ведёт к более высокому качеству классификации по сравнению с SimpleCNN. AdvancedCNN — оптимальный компромисс между производительностью и ресурсозатратами.

MPNN — самая передовая и эффективная из всех рассмотренных моделей. Используя данные, извлечённые с помощью MediaPipe, — компактный вектор из 63 признаков, описывающих ключевые точки кисти — модель значительно снижает размерность входных данных и фокусируется именно на информативных характеристиках жеста. Это обеспечивает лучшие результаты как по точности классификации, так и по скорости обучения. MPNN демонстрирует стабильную работу без переобучения, быстрое обучение и высокую общую производительность. В задачах перевода букв языка жестов она является наиболее перспективным решением, благодаря сочетанию высокой эффективности и качества распознавания.

# ЗАКЛЮЧЕНИЕ

В рамках курсового проекта была проведена разработка и исследование нескольких нейронных сетей для задачи перевода букв языка жестов на основе изображений кисти руки. В работе использовались модели различной сложности — от простых полносвязных сетей до сверточных и моделей, использующих компактные признаки, извлечённые с помощью MediaPipe.

Ключевые результаты экспериментов показали, что эффективность моделей напрямую зависит от их архитектуры и предварительной обработки данных. Использование сверточных слоёв существенно улучшает качество распознавания за счёт выделения пространственных признаков, а применение специализированного предпроцессинга с помощью MediaPipe позволяет значительно снизить размер входных данных и повысить точность и скорость обучения.

При работе с данными, представленными в виде компактных векторов ключевых точек кисти, модель MPNN показала наилучшие результаты по точности классификации и времени обучения, что подтверждает важность качественного выделения признаков для задач перевода жестов. Простые полносвязные сети оказались недостаточно эффективными для данной задачи и демонстрировали низкое качество распознавания.

Полученные результаты позволяют сделать вывод о том, что успешное решение задачи перевода букв языка жестов требует использования сверточных архитектур и специализированных методов предварительной обработки, обеспечивающих информативное и компактное представление входных данных. Перспективным направлением дальнейших исследований является интеграция методов глубокого обучения с анализом скелетных данных для повышения точности и устойчивости моделей в реальных условиях.

# СПИСОК ЛИТЕРАТУРЫ

1. MediaPipe Hands // Google Developers. — URL: https://developers.google.com/mediapipe/solutions/vision/hand\_tracking (дата обращения: 23.05.2025).
2. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [Электронный ресурс] // arXiv preprint arXiv:1409.1556. — 2014. — URL: [https://arxiv.org/abs/1409.1556](https://arxiv.org/abs/1409.1556" \t "_new) (дата обращения: 23.05.2025).
3. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition [Электронный ресурс] // arXiv preprint arXiv:1512.03385. — 2015. — URL: [https://arxiv.org/abs/1512.03385](https://arxiv.org/abs/1512.03385" \t "_new) (дата обращения: 23.05.2025).
4. Oord A. van den, Dieleman S., Zen H., et al. WaveNet: A Generative Model for Raw Audio [Электронный ресурс] // arXiv preprint arXiv:1609.03499. — 2016. — URL: [https://arxiv.org/abs/1609.03499](https://arxiv.org/abs/1609.03499" \t "_new) (дата обращения: 23.05.2025).
5. Wang J., Song Y., Huang Z. Gesture Recognition Based on CNN and LSTM // IEEE Access. — 2019. — Т. 7. — С. 189119–189129.
6. Jiang W., Chen X., Yang Z. Hand Gesture Recognition Using CNN with MediaPipe Skeleton Features // Proceedings of the 2022 IEEE International Conference on Image Processing (ICIP). — 2022. — С. 2345–2349.
7. Lecun Y., Bottou L., Bengio Y., Haffner P. Gradient-Based Learning Applied to Document Recognition // Proceedings of the IEEE. — 1998. — Т. 86, № 11. — С. 2278–2324.
8. Goodfellow I., Bengio Y., Courville A. Deep Learning. — MIT Press, 2016. — 800 с.
9. Kingma D.P., Ba J. Adam: A Method for Stochastic Optimization [Электронный ресурс] // arXiv preprint arXiv:1412.6980. — 2014. — URL: [https://arxiv.org/abs/1412.6980](https://arxiv.org/abs/1412.6980" \t "_new) (дата обращения: 23.05.2025).
10. Buliga V., Lunyov A. Neural Network Approaches for Sign Language Recognition // Journal of Intelligent & Fuzzy Systems. — 2021. — Т. 40, № 1. — С. 627–637.
11. Silver D., Huang A., Maddison C.J., et al. Mastering the Game of Go with Deep Neural Networks and Tree Search // Nature. — 2016. — Т. 529, № 7587. — С. 484–489.
12. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely Connected Convolutional Networks // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — С. 4700–4708.
13. Simonyan K., Zisserman A. Two-Stream Convolutional Networks for Action Recognition in Videos [Электронный ресурс] // arXiv preprint arXiv:1406.2199. — 2014. — URL: [https://arxiv.org/abs/1406.2199](https://arxiv.org/abs/1406.2199" \t "_new) (дата обращения: 23.05.2025).
14. Zhao H., Shi J., Qi X., Wang X., Jia J. Pyramid Scene Parsing Network // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). — 2017. — С. 2881–2890.
15. Medina J., Roa A., Díaz G. A Review of Hand Gesture Recognition Using Vision-Based Techniques // IEEE Access. — 2020. — Т. 8. — С. 175225–175243.
16. Liu J., Shah M. Learning Hand Pose Estimation from Depth Images via CNNs // Proceedings of the IEEE International Conference on Computer Vision Workshops. — 2017. — С. 1905–1913.

Министерство образования Республики Беларусь

Учреждение образования «Брестский государственный Технический университет»

Кафедра ИИТ

Приложение А  
«Текст программы»

**Выполнил**

студент 3 курса,

ФЭИС, группы ИИ-23

Макаревич Н. Р.

**Проверил**

Туз И. С.

Брест 2025

# Код программы предоставлен на GitHub по QR коду:

