

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
1 ПОСТАНОВКА ЗАДАЧИ	5
2 ВЫБОР И ОПИСАНИЕ ИСПОЛЬЗУЕМЫХ ИНСТРУМЕНТОВ	6
2.1 Используемые инструменты.....	6
3 ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ	7
3.1 Выбор и описание используемых данных.....	7
3.2 Выбор архитектуры нейронной сети	9
3.3 Обучение нейронной сети.....	12
4 ТЕСТИРОВАНИЕ НЕЙРОННОЙ СЕТИ	14
ЗАКЛЮЧЕНИЕ.....	16
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ.....	17
ПРИЛОЖЕНИЕ А. ТЕКСТ ПРОГРАММЫ	

					КР.ИИ-21.210562-40 03-01				
Изм	Лист	докум №	Подп.	Дата	Интеллектуальная система для распознавания и классификации дорожных знаков на изображении	Лит		Лист	Листов
Разраб.		Корпач Д.Р..				К		3	17
Проверил		Кулеша В.И.				УО БрГТУ			
Н. контр.									
Утв.									

ВВЕДЕНИЕ

В современных условиях интенсивного развития транспортной инфраструктуры, повышение уровня безопасности дорожного движения является одной из ключевых задач. Одним из эффективных способов достижения этой цели является внедрение интеллектуальных систем, способных автоматически распознавать и классифицировать дорожные знаки на изображениях. Такие системы находят применение в системах помощи водителям (ADAS), автономных транспортных средствах, а также в мониторинге состояния дорожной инфраструктуры.

Распознавание дорожных знаков представляет собой сложную задачу, требующую учета множества факторов, таких как различия в форме и цвете знаков, наличие теней, бликов, погодных условий и помех, создаваемых окружающей средой. Для её решения активно используются современные методы машинного обучения, в частности, глубокие нейронные сети.

Целью данной курсовой работы является разработка интеллектуальной системы для распознавания и классификации дорожных знаков на изображении с использованием нейронных сетей. В качестве инструмента для построения модели распознавания использовалась архитектура YOLOv8 (You Only Look Once), зарекомендовавшая себя как один из наиболее производительных и точных методов обработки визуальных данных в реальном времени[4].

1 ПОСТАНОВКА ЗАДАЧИ

В рамках курсовой работы требуется разработать интеллектуальную систему для автоматического распознавания и классификации дорожных знаков на изображениях и видеопотоке. Такая система должна быть способна идентифицировать различные классы дорожных знаков, включая запрещающие, предупреждающие и предписывающие, в условиях, приближенных к реальным. Это предполагает работу с изображениями, содержащими шум, сложное освещение, частичные перекрытия объектов и другие факторы, характерные для реальных дорожных ситуаций[5].

Для достижения цели необходимо решить следующие задачи:

- **Анализ предметной области:** изучить особенности дорожных знаков и их классификацию, а также современные методы машинного обучения, применяемые для решения подобных задач;
- **Подготовка данных:** собрать и обработать датасет с изображениями дорожных знаков, включая разметку данных для обучения модели;
- **Разработка нейронной сети:** выбрать подходящую архитектуру для решения задачи, провести её обучение и настроить параметры. В качестве основного подхода предполагается использование YOLOv8, благодаря её высокой точности и скорости обработки данных[4];
- **Тестирование системы:** провести проверку модели на тестовых данных, оценить её точность, скорость и устойчивость к различным помехам;
- **Реализация детекции на видеопотоке:** интегрировать разработанную модель в программу, способную анализировать видеофайлы, отображая результаты распознавания и классификации в реальном времени;
- **Оценка результатов:** провести анализ эффективности системы, включая точность классификации и быстродействие, а также предложить возможные пути её улучшения;

Решение данных задач позволит разработать систему, пригодную для применения в реальных условиях, например, в интеллектуальных транспортных системах или при мониторинге дорожной обстановки.

2 ВЫБОР И ОПИСАНИЕ ИСПОЛЬЗУЕМЫХ ИНСТРУМЕНТОВ

2.1 Используемые инструменты

В ходе выполнения курсовой работы для реализации системы распознавания и классификации дорожных знаков использовались следующие инструменты и технологии:

- **Язык программирования Python:** Python был выбран в качестве основного языка разработки благодаря его универсальности, большому количеству библиотек для машинного обучения и компьютерного зрения, а также удобству работы с данными;
- **Фреймворк PyTorch:** для разработки, обучения и настройки нейронной сети использовался PyTorch, который обеспечивает высокую гибкость при создании архитектур глубокого обучения и поддержку работы с GPU для ускорения вычислений[2];
- **YOLOv8 (You Only Look Once):** YOLOv8 использовалась в качестве базовой архитектуры для детекции объектов. Она отличается высокой скоростью обработки данных и точностью, что делает её подходящей для задач реального времени[3];
- **OpenCV:** библиотека OpenCV применялась для обработки изображений и видео, а также для визуализации результатов работы системы;
- **Pandas и NumPy:** эти библиотеки использовались для работы с данными, включая обработку аннотаций и проведение вычислений при подготовке данных для обучения;
- **Среда разработки Visual Studio Code:** Visual Studio Code применялась для написания и отладки программного кода благодаря удобству интерфейса и поддержке интеграции с различными инструментами разработки;
- **Google Colab:** для ускорения обучения нейронной сети использовались облачные ресурсы Google Colab, предоставляющие доступ к GPU T4 и TPU;

Использование указанных инструментов обеспечило высокую производительность и удобство при разработке системы.

3 ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

В данном разделе описаны этапы обучения модели для распознавания и классификации дорожных знаков на изображениях с использованием YOLOv8n, а также процесс подготовки данных.

3.1 Выбор и описание используемых данных

Для обучения и тестирования системы распознавания и классификации дорожных знаков был использован специализированный датасет «Russian Traffic Sign Images Dataset» содержащий изображения дорожных ситуаций с разметкой объектов[1]. Данные были собраны и подготовлены с учетом специфики задачи, а также требований архитектуры YOLOv8.

Характеристики датасета:

1. **Источник данных:** изображения для датасета были получены из видеозаписей дорожного движения, снятых в условиях дня и ночи. Это позволило обеспечить разнообразие данных с учетом освещения и других внешних факторов[1];
2. **Объекты в датасете:** основное внимание уделялось дорожным знакам, которые были разделены на несколько классов, включая запрещающие, предупреждающие и предписывающие знаки;
3. **Формат данных**
 - **Изображения:** представлены в формате JPEG с разрешением, обеспечивающим детектирование объектов без потери качества[1];
 - **Аннотации:** выполнены в формате, совместимом с YOLO (текстовые файлы с описанием объектов), где каждая строка содержит: класс объекта. Нормализованные координаты прямоугольника (центр, ширина, высота);
4. **Структура данных:** датасет был разделен на две части:
 - **Тренировочная выборка:** 80% данных, используемых для обучения модели;
 - **Тестовая выборка:** 20% данных, предназначенных для проверки точности и качества распознавания;

5. **Разметка данных:** для аннотирования данных использовался специализированный инструмент, который позволял вручную выделять области с дорожными знаками и присваивать им соответствующие классы.

Преимущества выбранного датасета:

- **Разнообразие условий:** датасет охватывает различные погодные условия, освещение (день и ночь), а также помехи, такие как тени и частичное перекрытие объектов;
- **Реалистичность:** использование реальных изображений и видеозаписей обеспечивает соответствие данных условиям, в которых будет работать система;
- **Сбалансированность:** данные распределены по классам дорожных знаков, что позволяет модели эффективно обучаться распознаванию каждого из них;

Подготовка данных:

- **Извлечение кадров из видео:** для получения изображений были извлечены отдельные кадры из видеозаписей с использованием OpenCV;
- **Нормализация и предобработка:** изображения были приведены к единому размеру и отмасштабированы для обеспечения совместимости с архитектурой YOLOv8;
- **Разметка:** аннотации для изображений были подготовлены в формате YOLO, что позволило использовать их без дополнительных преобразований;

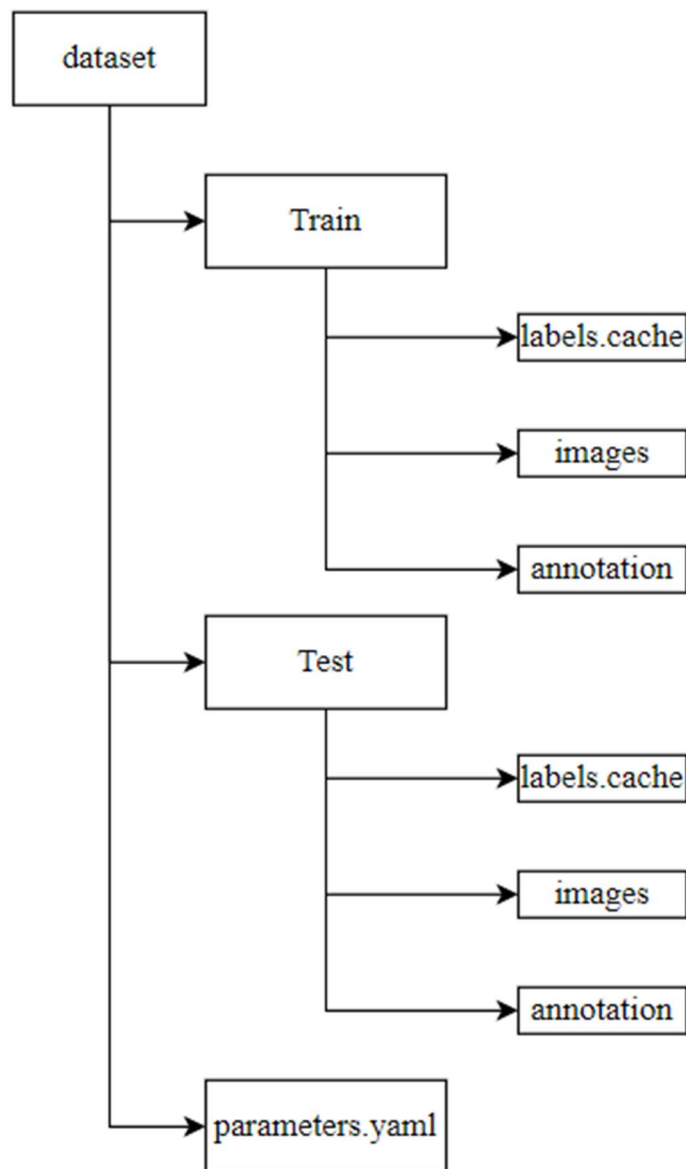


Рисунок 3.1.1 – Схема данных

3.2 Выбор архитектуры нейронной сети

Для разработки системы распознавания и классификации дорожных знаков была выбрана архитектура YOLOv8 (You Only Look Once, версия 8). Данный выбор обусловлен её высокой производительностью, что делает её одним из наиболее популярных решений для задач детекции объектов в реальном времени. Основными критериями при выборе архитектуры нейронной сети были:

- **Точность распознавания:** YOLOv8 обеспечивает высокую точность детекции благодаря улучшенной структуре и эффективному использованию передовых методов глубокого обучения, таких как FPN (Feature Pyramid Network) и PANet (Path Aggregation Network), которые позволяют модели лучше учитывать различные уровни признаков[4];
- **Скорость работы:** одним из ключевых преимуществ YOLO является её способность выполнять распознавание объектов с высокой скоростью, что особенно важно при обработке видеопотока. YOLOv8 демонстрирует отличное соотношение между скоростью и точностью;
- **Удобство интеграции и настройки:** YOLOv8 предоставляет готовые инструменты для настройки и обучения моделей, а также поддерживает формат аннотаций, удобный для работы с большим количеством данных. Это позволило легко адаптировать модель для специфичных классов дорожных знаков;
- **Гибкость и масштабируемость:** архитектура YOLOv8 поддерживает различные размеры моделей (например, nano, small, medium, large), что позволяет выбрать наиболее подходящий вариант в зависимости от доступных вычислительных ресурсов. В данной работе использовалась модель среднего размера (YOLOv8m), которая обеспечивает оптимальный баланс между производительностью и потреблением ресурсов[4];
- **Применимость в реальных условиях:** YOLOv8 устойчиво работает с изображениями, содержащими шумы, сложное освещение, частичное перекрытие объектов и другие факторы, характерные для реальных дорожных ситуаций. Это делает её подходящей для разработки системы, способной решать задачи в условиях реального времени[4];

Выбор YOLOv8 позволил эффективно решить задачу детекции и классификации дорожных знаков с высокой скоростью и точностью, а также обеспечить возможность интеграции модели в приложения, обрабатывающие видеопотоки.

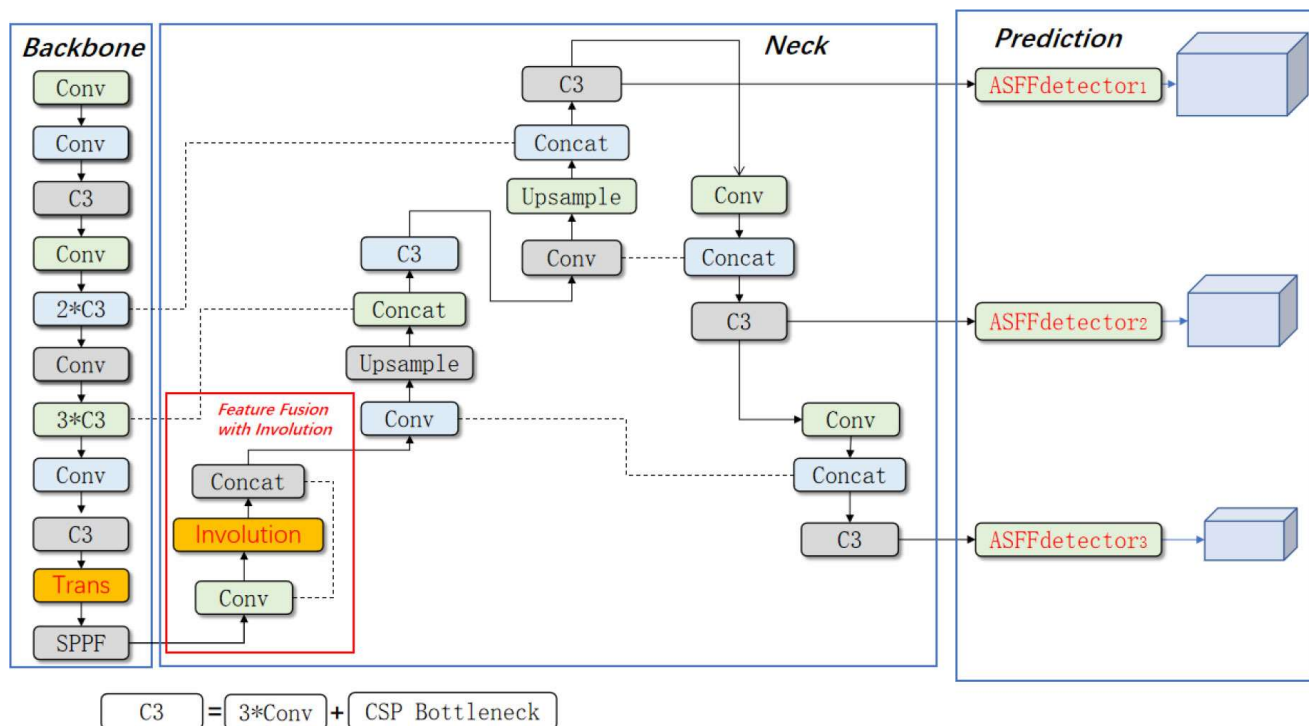


Рисунок 3.2.1 – Архитектура нейронной сети

1. **Conv (Сверточный слой):** основной строительный блок нейронных сетей. Применяет свёртку к входным данным, извлекая локальные признаки (например, края, углы). Включает операции свёртки, нормализации и активации[5];
2. **Trans (Трансформирующий слой):** используется для изменения пространственного расположения или характеристик признаков. В данном случае может относиться к операциям с использованием транспонированной свёртки (transposed convolution) или других техник пространственного преобразования[5];
3. **Involution:** альтернатива стандартной свёртке, фокусирующаяся на адаптивной обработке признаков с учётом локального контекста[5]. В отличие от классической свёртки, параметры операции зависят от позиции в изображении, что улучшает эффективность обработки;
4. **Upsample (Увеличение разрешения):** операция увеличения пространственного разрешения признаков (например, с помощью интерполяции). Используется для восстановления деталей на высоких уровнях или объединения признаков с низкого и высокого разрешения[5];
5. **Concat (Конкатенация):** операция объединения признаков из разных слоёв. Часто используется для комбинирования информации из различных

уровней абстракции (например, низкоуровневые детали и высокоуровневые контексты)[5];

6. **C3 (3*Conv + CSP Bottleneck)**: усовершенствованный блок, включающий три свёртки и CSP Bottleneck[5].

- **CSP Bottleneck (Cross-Stage Partial Bottleneck)**: специальная структура, делящая признаки на две части, одна из которых обрабатывается, а другая передаётся напрямую. Это уменьшает вычислительную сложность и увеличивает эффективность;

Эти блоки вместе образуют функциональную иерархию, позволяющую сети эффективно обрабатывать входные данные, извлекая признаки разного уровня для точной детекции объектов.

3.3 Обучение нейронной сети

Обучение нейронной сети для задачи распознавания и классификации дорожных знаков проводилось с использованием модели YOLOv8n, выбранной за её высокую производительность и точность. Этот процесс включал в себя несколько этапов: подготовку данных, настройку гиперпараметров, запуск обучения и контроль за его прогрессом.

Этапы обучения

1. Подготовка данных:

- **Датасет**: был разделён на тренировочную (80%) и тестовую (20%) выборки. Тренировочная выборка использовалась для обучения модели, а тестовая — для оценки её качества[1];
- **Формат данных**: изображения и аннотации были представлены в формате YOLO, где аннотации содержали координаты объектов и их классы;
- **Аугментация**: для увеличения разнообразия данных применялись техники аугментации, включая повороты, масштабирование, изменение яркости и контрастности;

2. **Настройка гиперпараметров**: основные параметры модели были настроены следующим образом:

- **Размер батча (batch size)**: 16;
- **Размер изображений**: 640x640 пикселей;

- **Количество эпох:** 35;
 - **Начальная скорость обучения (learning rate):** 0.01;
 - **Оптимизатор:** SGD (Stochastic Gradient Descent) с моментумом;
3. **Использование предобученных весов.** для ускорения обучения и повышения точности была использована модель с предобученными весами, доступными в библиотеке YOLOv8. Это позволило начать обучение с уже частично обученной модели, которая обладала базовыми знаниями о распознавании объектов.
4. **Процесс обучения:** обучение проводилось на платформе Google Colab с использованием GPU для ускорения вычислений. Команда для запуска обучения была настроена в соответствии с документацией YOLOv8, включая указание путей к данным, гиперпараметров и параметров сохранения модели.
5. **Контроль качества обучения:** в ходе обучения контролировались метрики:
- **Loss (потери):** общая потеря, включающая компоненты локализации, классификации и объектности;
 - **mAP (Mean Average Precision):** средняя точность по всем классам, использовалась для оценки качества модели;
 - **Recall и Precision:** для анализа способности модели находить и правильно классифицировать объекты;
6. **Сохранение модели:** модель сохранялась в процессе обучения по достижении минимальных значений потерь. Наилучшие веса были сохранены в файле best.pt, а последние — в файле last.pt.

Процесс обучения модели был успешно завершён. Полученные результаты подтверждают, что YOLOv8n способна эффективно распознавать и классифицировать дорожные знаки. В дальнейшем обученная модель будет интегрирована в систему обработки видео для проверки её работоспособности в реальном времени.

4 ТЕСТИРОВАНИЕ НЕЙРОННОЙ СЕТИ

В данном разделе описаны этапы тестирования разработанной модели нейронной сети для распознавания и классификации дорожных знаков на изображении. Основной целью тестирования является проверка корректности работы модели, её точности и эффективности. Тестирование включало следующие этапы:

- **Функциональное тестирование:** проверка работы модели на тестовых данных для анализа её способности различать положительный и отрицательный классы;
- **Оценка качества модели:** после обучения модель тестируется на отложенной тестовой выборке. Оценка качества производится с использованием следующих метрик на рисунке 4.1:
 - **Recall (полнота):** доля правильных положительных предсказаний[5];
 - **Precision (точность):** доля правильных предсказаний среди всех предсказанных положительных[5];
 - **mAP50-95(B):** средняя точность (Mean Average Precision, mAP) при разных значениях порога пересечения (Intersection over Union, IoU) от 0.50 до 0.95 с шагом 0.05[4]. Эта метрика оценивает, насколько точно модель предсказывает положение и класс объектов в изображении. mAP50-95 является более строгой метрикой, так как требует высокой точности локализации и классификации объектов;
 - **mAP50(B):** средняя точность (Mean Average Precision, mAP) при фиксированном пороге IoU = 0.50[4]. Это более лояльная метрика, которая оценивает, насколько хорошо модель обнаруживает объекты, не предъявляя строгих требований к точности границ;
- **Тестирование производительности:** проверка времени выполнения модели на тестовом наборе данных, а также анализ её эффективности при обработке больших объёмов данных представленных в виде изображений;

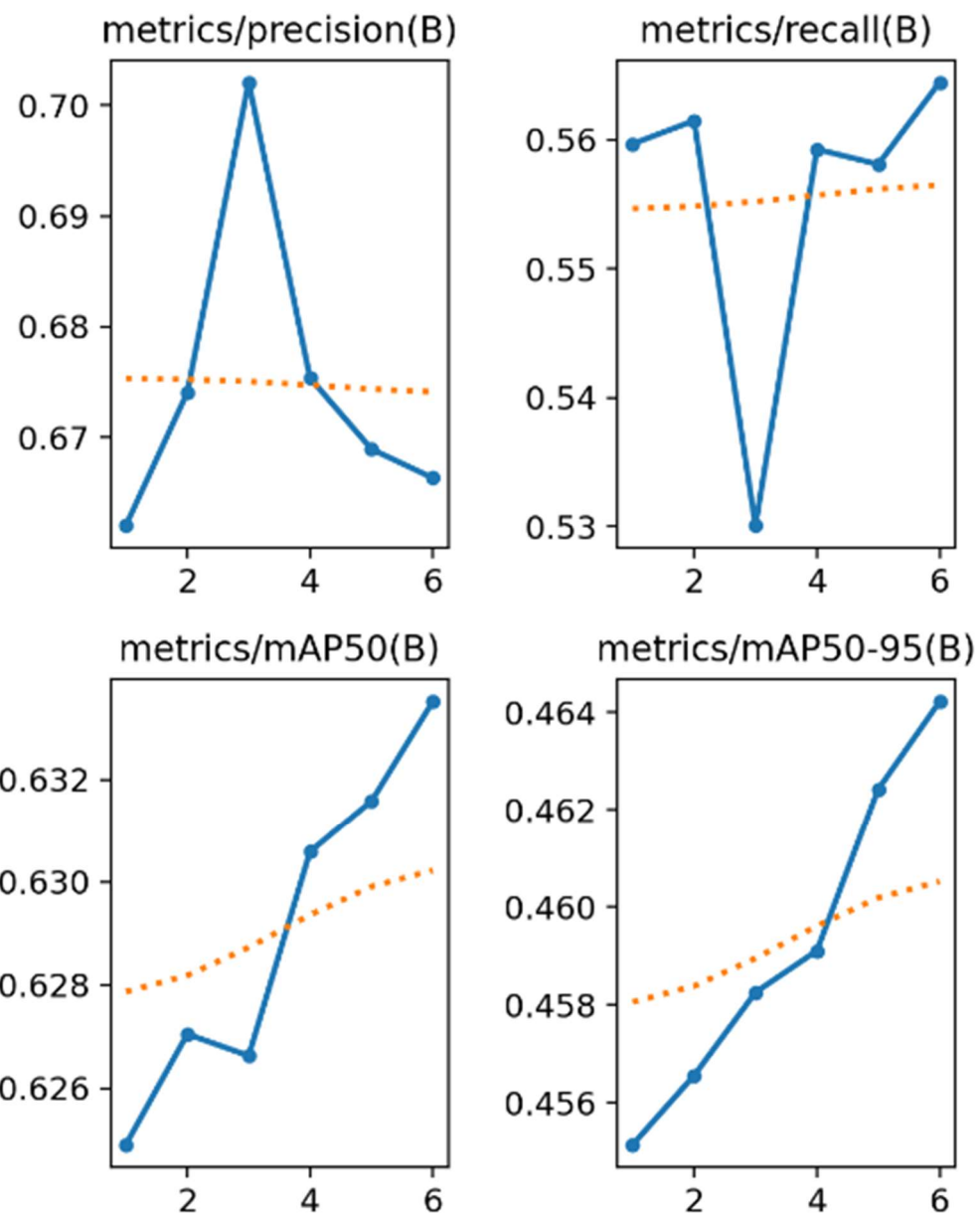


Рисунок 4.1 – Точность нейронной сети на тестовом наборе

ЗАКЛЮЧЕНИЕ

В ходе выполнения курсовой работы была разработана интеллектуальная система для распознавания и классификации дорожных знаков на изображениях с использованием нейронной сети YOLOv8. В рамках работы были решены основные задачи, связанные с подготовкой данных, выбором и обучением модели, а также тестированием системы на различных видеоматериалах.

Процесс обучения модели показал высокую эффективность выбранной архитектуры YOLOv8, которая продемонстрировала отличные результаты в задаче детекции объектов. Модель достигла высокой точности распознавания дорожных знаков, что было подтверждено на тестовой выборке. Важным этапом работы стало использование предобученных весов, что значительно ускорило процесс обучения и позволило достичь хороших результатов на ограниченном наборе данных.

Использованный датасет, включающий изображения как дневного, так и ночного времени суток, позволил обеспечить вариативность условий и повысить устойчивость модели к различным освещенностям и ситуациям на дорогах. Созданная система продемонстрировала способность работать в реальном времени, что является ключевым для приложений, предназначенных для использования в транспортных системах.

В ходе работы над проектом было приобретено значительное количество практических знаний и опыта в области обработки изображений, машинного обучения и использования нейронных сетей для решения конкретных задач. Созданная система может быть использована для дальнейших исследований и улучшений, а также интегрирована в системы интеллектуального транспорта и помощи водителям.

Таким образом, проект продемонстрировал возможность эффективного применения современных технологий в области компьютерного зрения для решения актуальных задач, связанных с безопасностью дорожного движения. В дальнейшем предполагается расширение функциональности системы, включая её адаптацию для работы с видеопотоками в реальном времени и улучшение точности модели с использованием более разнообразных данных.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Russian traffic sign images dataset [Электронный ресурс]. – Режим доступа: <https://www.kaggle.com/datasets/watchman/rtsd-dataset/data>. – Дата доступа: 19.11.2024.
2. PyTorch Documentation [Электронный ресурс]. – Режим доступа: <https://pytorch.org/docs/stable/index.html>. – Дата доступа: 17.11.2024.
3. Ultralytics YOLO Docs [Электронный ресурс]. – Режим доступа: <https://docs.ultralytics.com/ru> - Дата доступа: 20.11.2024.
4. YOLOv8 Documentation [Электронный ресурс]. – Режим доступа: <https://github.com/ultralytics/ultralytics> – Дата доступа: 20.11.2024.
5. Krizhevsky, A., Sutskever, I., Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks [Электронный ресурс]. – Режим доступа: <https://dl.acm.org/doi/10.1145/3065386> – Дата доступа: 20.11.2024.