

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»**

Кафедра _____ математического анализа _____

Нейронные сети

в задаче распознавания дорожных знаков

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента _____ 2 _____ курса _____ 218 _____ группы
направление _____ 01.04.02 — Прикладная математика и информатика _____

_____ механико-математического факультета _____

_____ Тюрина Кирилла Романовича _____

Научный руководитель
доцент, к.т.н., доцент _____ А.К. Кузьмин _____

Заведующий кафедрой
зав.кафедрой, к.ф.-м.н., доцент _____ Е.В. Разумовская _____

Саратов 2026

Введение. Системы автоматического распознавания дорожных знаков представляют собой важный инструмент обеспечения безопасности дорожного движения и являются одним из ключевых компонентов интеллектуальных транспортных систем. Их применение способствует снижению числа ошибок со стороны водителей, повышению надежности функционирования транспортных средств, а также улучшению организации дорожного движения в условиях интенсивного городского трафика.

Современное развитие транспортной отрасли характеризуется активной цифровизацией и автоматизацией, что повышает требования к системам восприятия и обработки информации. В этих условиях особое значение приобретают технологии, позволяющие эффективно интерпретировать дорожные знаки при различных условиях освещенности, погодных явлениях и при наличии помех. Одним из наиболее результативных инструментов решения данной задачи являются сверточные нейронные сети (CNN), зарекомендовавшие себя как надежный метод анализа и классификации изображений. Их применение в системах помощи водителю и проектах автономного вождения подтверждает высокую актуальность данного направления исследований.

До широкого применения глубокого обучения основные подходы к распознаванию дорожных знаков строились на анализе цвета и формы в различных цветовых пространствах (RGB, HSV, YCbCr), методах HOG (Histogram of Oriented Gradients), SIFT (Scale-Invariant Feature Transform), Gabor-фильтрах, каскадах Хаара с SVM и алгоритме k -ближайших соседей. Все перечисленные методы требуют ручного выбора признаков и плохо масштабируются на изображения, полученные в сложных погодных условиях. Сверточные нейронные сети позволили кардинально повысить точность классификации благодаря способности автоматически извлекать иерархические признаки непосредственно из изображений без необходимости определять признаки вручную.

Целью работы является обучение эффективной сверточной нейросетевой модели распознавания российских дорожных знаков и ее внедрение в программный продукт.

Для достижения поставленной цели решены следующие задачи:

- проведен анализ существующих методов распознавания дорожных знаков, включая классические подходы на основе ручных признаков и современные методы глубокого обучения;
- выполнен обзор наиболее распространенных архитектур сверточных нейронных сетей и проведен сравнительный анализ их характеристик;
- обосновано использование специализированной архитектуры CNN, разработанной с учетом особенностей входных данных, и описана ее структура;
- подготовлено обучающее множество на основе датасета RTSD, включая предварительную обработку и аугментацию данных;
- спроектирована модульная архитектура программного решения, реализующего полный конвейер обнаружения и классификации дорожных знаков;
- разработаны основные модули программного продукта: модуль загрузки и предварительной обработки данных, модуль обнаружения дорожных знаков, модуль классификации и модуль пользовательского интерфейса;
- определены оптимальные макропараметры модели и настройки обучения посредством экспериментального исследования;
- проведено обучение двух версий модели – исходной и улучшенной архитектуры – с последующим сравнительным анализом полученных результатов;
- выполнена оценка качества обученной модели с использованием метрик точности, полноты, F_1 -меры и матрицы ошибок на тестовой выборке;
- осуществлено внедрение обученной модели в программный продукт с применением формата сохранения HDF5;
- проведено функциональное тестирование программного продукта по пяти сценариям, выявленные дефекты устранены;
- дана оценка эксплуатационных возможностей разработанной системы на репрезентативной тестовой выборке из 80 изображений;
- подготовлена техническая документация, описывающая архитектуру системы, алгоритмы ее компонентов и результаты тестирования.

Структура магистерской работы. В магистерской работе содержится введение, 9 разделов и заключение.

В первом разделе проведен сравнительный анализ методов распознавания дорожных знаков: от классических алгоритмов компьютерного зрения до современных архитектур на основе глубокого обучения. Рассмотрены метод HOG, алгоритм SIFT, Gabor-фильтры, каскады Хаара с AdaBoost, алгоритм k -NN. Среди нейросетевых подходов проанализированы одноэтапные детекторы (YOLO, SSD, Faster R-CNN), двухэтапные методы на основе FCN, а также гибридные архитектуры CNN+ViT и легкие сети IECES. Составлена сравнительная таблица методов с описанием принципов работы, преимуществ и недостатков каждого подхода.

Во втором разделе представлен обзор наиболее распространенных архитектур CNN: LeNet-5, AlexNet, VGGNet (VGG-16, VGG-19), GoogLeNet, ResNet, DenseNet, MobileNet и EfficientNet. Описана эволюция архитектур от компактных сетей (LeNet-5, 60 тыс. параметров) до сверхглубоких и высокоточных моделей (ResNet-152), а также энергоэффективных решений для мобильных платформ с применением глубинно-разделимых сверток (MobileNet) и принципа compound scaling (EfficientNet).

В третьем разделе обоснован выбор собственной архитектуры CNN для решения задачи классификации дорожных знаков датасета RTSD. Приведено описание и детальный сравнительный анализ исходной и улучшенной архитектур модели, включая графики обучения, матрицу ошибок и метрики качества на тестовой выборке.

В четвертом разделе изложена теоретическая база сверточных нейронных сетей. Рассмотрены основные строительные блоки CNN: сверточный слой, функции активации ReLU и Softmax, слой подвыборки (MaxPooling, AvgPooling), пакетная нормализация, полносвязный слой, регуляризация Dropout. Описан алгоритм обратного распространения ошибки и оптимизатор Adam с первым и вторым моментами градиента.

В пятом разделе описан набор данных RTSD (Russian Traffic Sign Images Dataset). Для обучения выбран датасет rtsd-r3, содержащий 93 218 изображений российских дорожных знаков, распределенных по 96 классам: 70 293

в обучающем наборе и 22 925 – в тестовом. Представлены примеры изображений из каждого класса и распределение изображений по классам.

В шестом разделе представлено подробное описание архитектур исходной и улучшенной моделей с программным кодом и анализом результатов обучения. Проведен анализ классов с наименьшей точностью классификации.

В седьмом разделе описаны основные модули разработанного программного продукта: модуль загрузки и предобработки данных, модуль обнаружения знаков на основе HSV-сегментации, модуль классификации (CNN) и модуль графического интерфейса (PyQt6).

В восьмом разделе представлены результаты тестирования программного продукта по пяти сценариям, описаны выявленные дефекты детектора и способы их устранения.

В девятом разделе проведена оценка эксплуатационных возможностей программного продукта. На основе результатов системного тестирования проанализированы качество обнаружения дорожных знаков и точность их классификации, а также рассмотрены нефункциональные характеристики системы, включая время обработки изображений, устойчивость к исключительным ситуациям и совместимость с операционной системой. Кроме того, выполнено сравнение полученных результатов с аналогичными решениями, использующими методы распознавания дорожных знаков на датасете RTSD, и сформулированы выявленные ограничения разработанной системы.

Основное содержание работы. Рассматривается задача распознавания российских дорожных знаков с применением двухкомпонентной системы: HSV-детектора для локализации областей интереса и сверточной нейронной сети для их классификации.

Набор данных. Для обучения модели использован набор изображений дорожных знаков RTSD. Пример изображений из каждого класса представлен в соответствии с рисунком 1; распределение изображений по 50 наиболее часто встречающимся классам приведено в соответствии с рисунком 2.



Рисунок 1 — Примеры изображений из каждого класса датасета rtsd-r3

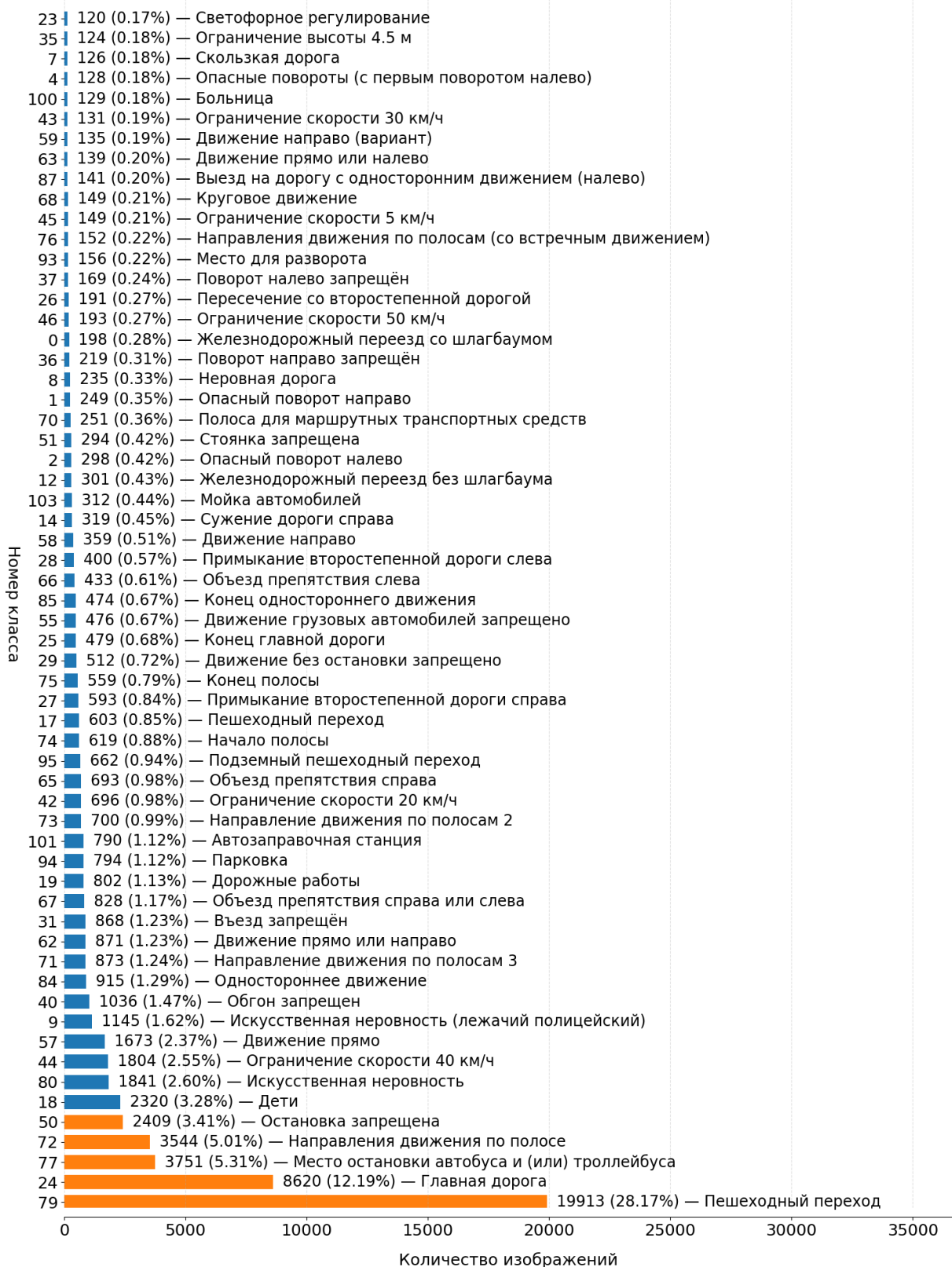


Рисунок 2 — Распределение изображений по наиболее распространенным классам

Архитектура модели. Для решения задачи разработана собственная архитектура сверточной нейронной сети, учитывающая специфику входных изображений малого разрешения (48×48 пикселей). Улучшенная модель включает четыре сверточных блока с постепенным наращиванием числа фильтров ($32 \rightarrow 64 \rightarrow 128 \rightarrow 256$), BatchNormalization после каждого сверточного слоя, GlobalAveragePooling2D и прогрессивную стратегию Dropout ($0.25 \rightarrow 0.3 \rightarrow 0.4 \rightarrow 0.5$).

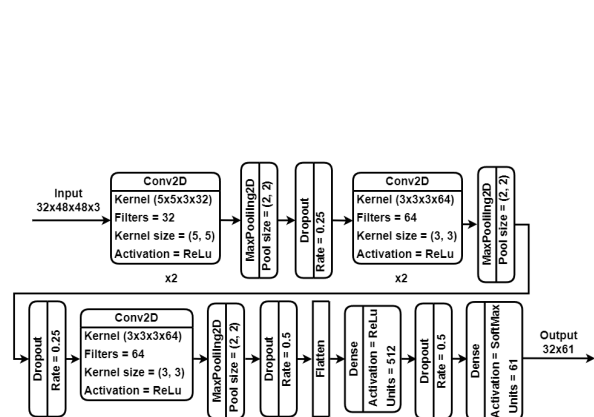


Рисунок 3 — Исходная архитектура модели

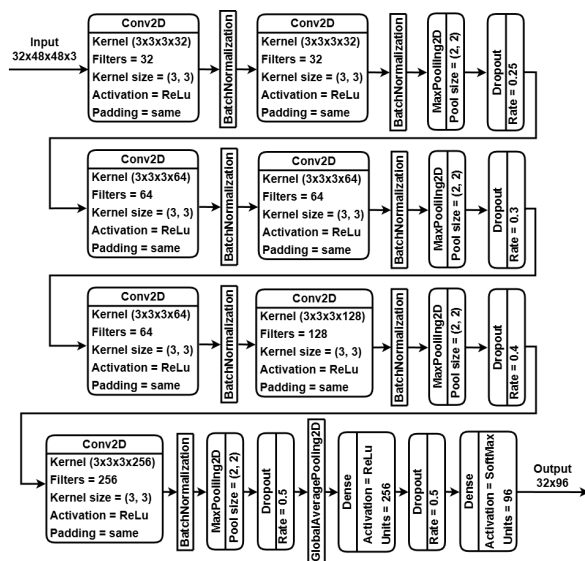


Рисунок 4 — Улучшенная архитектура модели

Общее число обучаемых параметров улучшенной модели составляет 674 048, из которых лишь $\approx 13.4\%$ приходится на полносвязную часть (в исходной архитектуре полносвязная часть составляла $\approx 80.3\%$ от общего числа параметров модели).

В процессе обучения применялся оптимизатор Adam с начальной скоростью обучения $\eta = 0.0005$, механизм ранней остановки (EarlyStopping) и адаптивного снижения скорости обучения (ReduceLROnPlateau). В качестве функции потерь использовалась категориальная кросс-энтропия:

$$L = - \sum_{c=1}^C y_c \log \hat{y}_c,$$

где C – число классов, y_c – истинная метка в формате one-hot, \hat{y}_c – предсказанная вероятность класса c .

Динамика обучения. Графики точности и функции потерь улучшенной модели приведены в соответствии с рисунком 5, с рисунком 6.

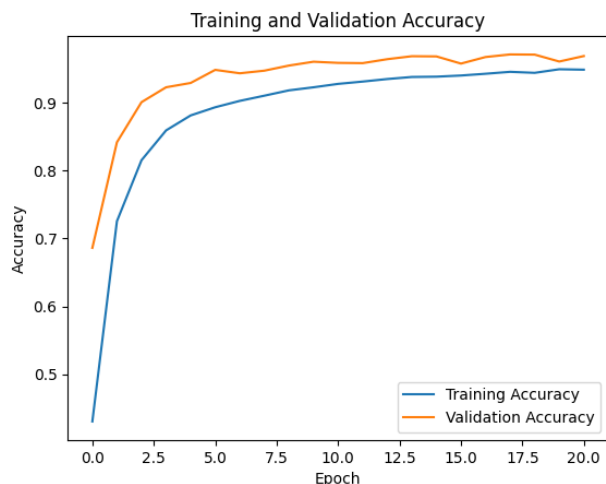


Рисунок 5 — График точности улучшенной модели

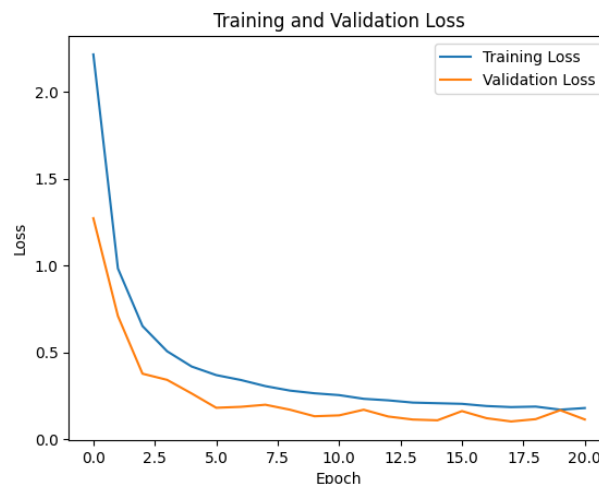


Рисунок 6 — График потерь улучшенной модели

Улучшенная модель достигла значения $val_accuracy = 98.28\%$ к 31-й эпохе, превысив итоговый результат исходной модели (97.48%) на 0.80 процентных пункта.

Оценка качества. Метрики качества моделей на тестовой выборке представлены в соответствии с таблицей 1.

Таблица 1 — Метрики качества моделей на тестовой выборке

Метрика	Исходная модель	Улучшенная модель
Accuracy	97.00%	98.32%
Precision (weighted)	97.40%	98.49%
Recall (weighted)	97.00%	98.32%
F_1 -мера (weighted)	96.85%	98.25%

Матрица ошибок улучшенной модели приведена в соответствии с рисунком 7. Выраженная диагональная структура свидетельствует о высоком качестве классификации по большинству из 96 классов.

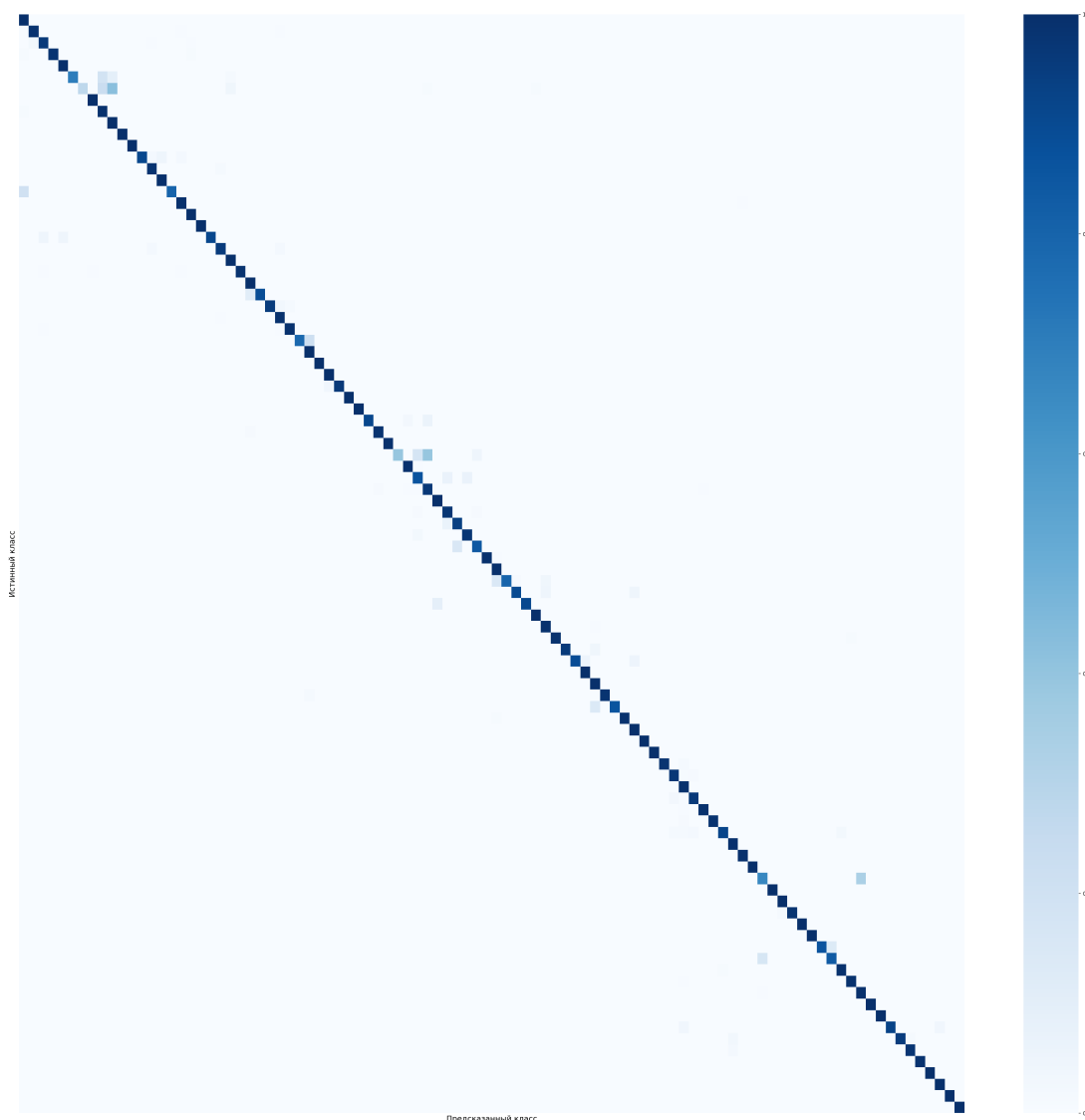


Рисунок 7 — Матрица ошибок улучшенной модели

Модуль обнаружения. Для локализации потенциальных областей дорожных знаков выполняется сегментация в пространстве HSV с формированием итоговой бинарной маски:

$$M = M_{\text{red}} \cup M_{\text{blue}} \cup M_{\text{yellow}}.$$

К маске применяются морфологическое открытие (устранение шумов) и закрытие (заполнение разрывов) с ядром 5×5 . Кандидаты фильтруются по площади (≥ 1500 пкс), коэффициенту заполненности ($S \geq 0.6$) и соотношению сторон ($w/h \in (0.5; 1.8)$). Финальная обработка выполняется алгорит-

мом Non-Maximum Suppression с порогом $\text{IoU} = 0.4$:

$$\text{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|}.$$

Программный продукт. Разработанный программный продукт реализует полный конвейер обработки изображений. Интерфейс приложения представлен в соответствии с рисунком 8; примеры распознавания дорожных знаков – в соответствии с рисунком 9.

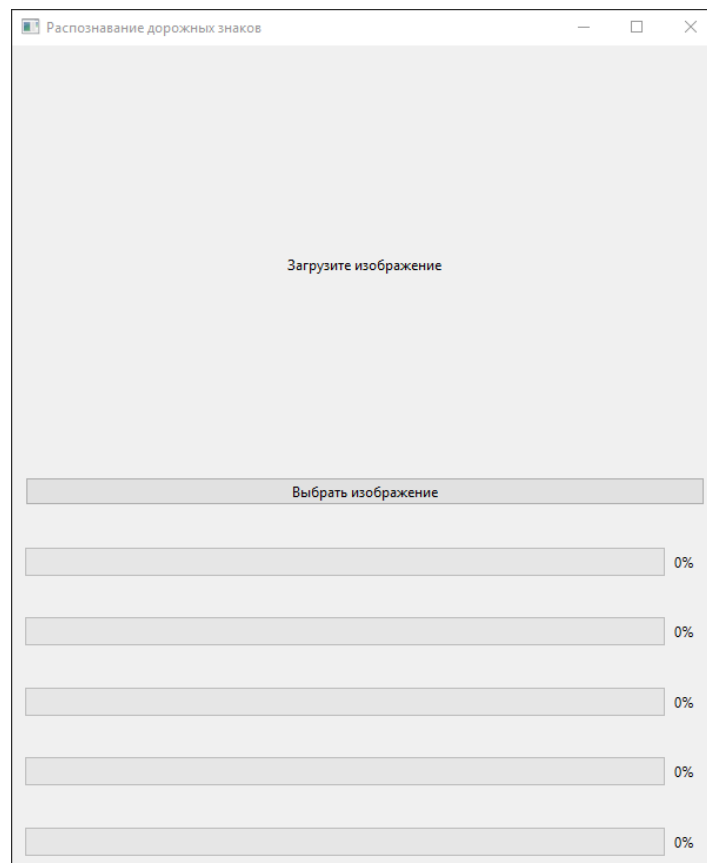


Рисунок 8 — Интерфейс программного продукта

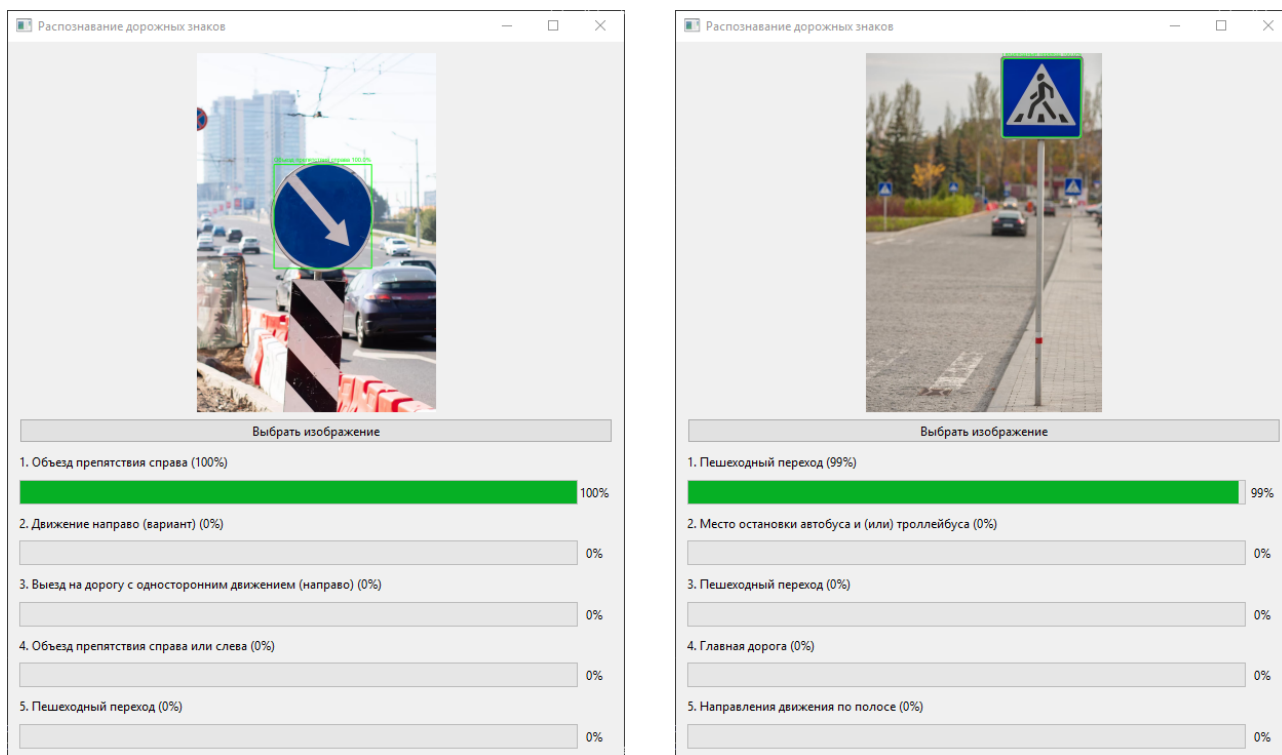


Рисунок 9 — Примеры распознавания дорожных знаков

Результаты тестирования. Модуль обнаружения оценивался на тестовой выборке из 80 изображений, полученных в реальных дорожных условиях. Результаты оценки модуля обнаружения представлены в соответствии с таблицей 2.

Таблица 2 — Метрики качества модуля обнаружения дорожных знаков

Метрика	Значение
Точность (Precision)	1.000
Полнота (Recall)	0.975
F_1 -мера	0.987
Среднее время обработки	0.53 с
Ложноположительные срабатывания	0 из 80

По результатам сравнения с аналогичными решениями на датасете RTSD разработанная система достигла $\text{Recall} = 0.975$ – против 0.900 у базовой HSV+CNN-системы и 0.890 у подхода с HOG+SVM в качестве детектора (Shakhuro et al., 2016).

Примеры работы системы при тестировании по сценариям с одним знаком, несколькими знаками и при частичном перекрытии представлены в соответствии с рисунком 10.

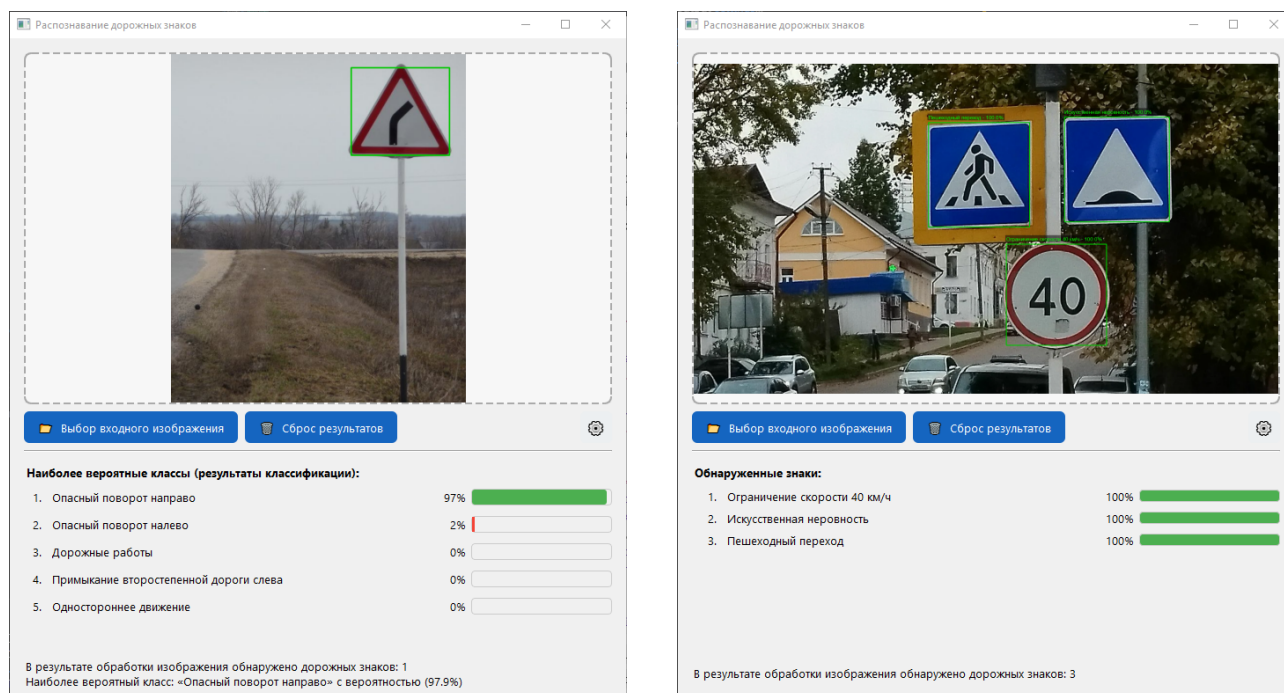


Рисунок 10 — Результаты тестирования: один знак (слева), несколько знаков (справа)

Заключение. В рамках магистерской работы проведено исследование методов автоматического распознавания дорожных знаков на основе сверточных нейронных сетей. Разработана, обучена и внедрена в программный продукт собственная модель классификации российских дорожных знаков. Решены все задачи, поставленные во введении.

Проанализированы современные подходы к построению CNN для задач классификации изображений. На их основе разработана архитектура модели, адаптированная к особенностям датасета RTSD. В ходе последовательного совершенствования архитектуры определены оптимальные гиперпараметры и настройки обучения – нормализация по батчам, механизмы регуляризации и адаптивное изменение скорости обучения, – что позволило повысить устойчивость к переобучению и улучшить обобщающую способность сети.

Разработанный программный продукт реализует полный конвейер обработки: от локализации областей интереса посредством цветовой HSV-

сегментации с многоуровневой фильтрацией до классификации обнаруженных областей нейросетевой моделью и визуализации результатов в графическом интерфейсе. Модульная архитектура продукта обеспечивает независимость компонентов и возможность их замены и масштабирования.

Практическая значимость полученных результатов определяется их направленностью на задачи обеспечения безопасности дорожного движения. Разработанное решение может быть использовано в системах помощи водителю и элементах автономного управления транспортными средствами. Перспективными направлениями дальнейших исследований являются замена HSV-детектора нейросетевой моделью детекции для повышения работоспособности в сложных условиях съемки, а также оптимизация модели для развертывания на встраиваемых платформах с ограниченными ресурсами.