

УДК 004.932.2:629.34
DOI 10.17513/snt.40604

ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ УРОВНЯ ЗАГРУЖЕННОСТИ ПАССАЖИРСКОГО ТРАНСПОРТА НА ОСНОВЕ АНАЛИЗА ВИДЕОДАННЫХ ИЗ САЛОНА

**Долженкова М.Л., Колупаев А.В. ORCID ID 0009-0008-0656-4479,
Метелев А.П. ORCID ID 0000-0001-7246-9156,
Ржаникова Е.Д. ORCID ID 0000-0002-3266-4747**

*Федеральное государственное бюджетное образовательное учреждение высшего образования
«Вятский государственный университет», Киров, Российская Федерация,
e-mail: ap_metelev@vyatsu.ru*

В работе рассматривается проблема определения уровня загруженности городского пассажирского транспорта и предлагается метод автоматической оценки пассажиропотока в режиме реального времени, основанный на применении технологий искусственного интеллекта. Предметом исследования является методология оценки загруженности общественного транспорта на основе оперативного анализа видеоданных из салона автобуса. Целью исследования является разработка системы, которая при помощи технологий машинного обучения определяет загруженность единиц общественного транспорта, оснащенных видеосъемками в салоне. Методология проведения исследований включала разработку и последующее обучение нейронной сети. Разработана полносвязная многослойная нейронная сеть для классификации изображений, обучение которой производилось на основе тренировочных выборок видеоданных, полученных из салонов автобусов, с экспертными оценками загруженности по пятибалльной шкале. В качестве функции потерь использовалась кросс-энтропия, а для оптимизации процесса машинного обучения применялся алгоритм Adam. Обучение сети проводилось с перемешиванием тренировочной выборки и отсевом нейронов на каждой итерации. Средняя вероятность верной классификации изображений на тестовой выборке составила 88 %. Модель продемонстрировала высокую точность классификации, быстроедействие и невысокие требования к аппаратным средствам. Разработанная система может быть использована для информирования ожидающих транспорт пассажиров, повышения качества организации движения городского транспорта и анализа эффективности пассажирских перевозок. Статья будет полезна специалистам в сфере городского транспорта, разработчикам систем управления транспортной инфраструктурой, исследователям в области машинного обучения и искусственного интеллекта.

Ключевые слова: нейронная сеть, машинное обучение, искусственный интеллект, общественный пассажирский транспорт, пассажиропоток, анализ видео

MACHINE LEARNING-BASED APPROACH FOR DETERMINING PUBLIC TRANSPORT OCCUPANCY USING INTERIOR VIDEO ANALYSIS

**Dolzhenkova M.L., Kolupaev A.V. ORCID ID 0009-0008-0656-4479,
Metelev A.P. ORCID ID 0000-0001-7246-9156,
Rzhanikova E.D. ORCID ID 0000-0002-3266-4747**

*Federal State Budgetary Educational Institution of Higher Education "Vyatka State University",
Kirov, Russian Federation, e-mail: ap_metelev@vyatsu.ru*

This paper considers the problem of determining the level of occupancy in urban passenger transport and proposes a method for automatic assessment of passenger flow in real time, based on the application of artificial intelligence technologies. The subject of the study is a methodology for assessing public transport occupancy based on the operational analysis of video data from inside the bus. The aim of the work is to develop a system that uses machine learning technologies to determine the occupancy of public transport units equipped with video cameras in the cabin. The research methodology included the development and subsequent training of a neural network. A fully connected multilayer neural network for image classification was developed, which was trained on the basis of training samples of video data obtained from inside the buses, with expert assessments of congestion on a five-point scale. Cross-entropy was used as the loss function, and the Adam algorithm was applied to optimize the machine learning process. Network training was carried out by shuffling the training sample and dropout of neurons at each iteration. The average probability of correct classification of images on the test samples was 88 %. The model demonstrated high classification accuracy, fast performance, and low hardware requirements. The developed system can be used to inform passengers waiting for transport, improve the organization of urban transport, and analyze the efficiency of passenger transportation. This article will be useful for urban transportation specialists, developers of transport infrastructure management systems, and researchers in the fields of machine learning and artificial intelligence.

Keywords: neural network, machine learning, artificial intelligence, public passenger transport, passenger flow, video analysis

Введение

В годы цифровизации системы информирования пассажиров претерпели значительные изменения. Современные информационные технологии позволяют предоставлять пассажирам городского общественного транспорта актуальную информацию о расписании, маршрутах и местоположении транспорта в режиме реального времени, могут оповещать о прогнозируемом времени прибытия на остановку (рис. 1). Интеграция с мобильными приложениями и веб-сервисами (рис. 2) дает пассажирам возможность планировать поездки с учетом текущих условий на дорогах. Тем не менее большинство существующих систем информирования имеют существенный недостаток – отсутствие данных о загруженности транспортных средств в реальном масштабе времени. Эта проблема порождает ряд негативных последствий для всех участников транспортного процесса.

Для пассажиров отсутствие информации о загруженности транспорта означает потерю возможности адекватно оценить комфортность предстоящей поездки. Пассажиры лишены шанса выбрать наименее загруженный маршрут или время для поездки, что нередко приводит к неприятным ситуациям в переполненных салонах. В результате удовлетворенность транспортной системой снижается, и часть пассажиров может отказаться от использования общественного транспорта в пользу личных автомобилей или других видов передвижения.

Транспортные компании также несут существенные потери из-за отсутствия данных о загруженности. Не имея актуальных сведений, они сталкиваются с неэффективным планированием маршрутов и расписания движения, что приводит к недозагрузке или перегрузке транспортных средств. Возникают проблемы с рациональным использованием подвижного состава и ресурсов: компании могут выделять больше транспорта, чем необходимо, или, наоборот, не справляться с пассажиропотоком [1]. Это негативно сказывается на качестве обслуживания, ведет к потерям доходов и затрудняет оценку эффективности работы транспортной системы.

Кроме того, недостаток информации о загруженности транспорта осложняет решение более широких задач долгосрочного развития общественного транспорта, оптимизации маршрутной сети и дорожной инфраструктуры. Без точных данных о распределении пассажиропотока трудно определить, где и какие изменения необходимы, чтобы сделать транспортную систему более эффективной и удобной для пользователей.



Рис. 1. Информационное табло на остановке общественного транспорта
Примечание: составлен авторами



Рис. 2. Сайт центральной диспетчерской службы
Примечание: составлен авторами

Таким образом, проблема отсутствия данных о загруженности транспортных средств является одним из ключевых недостатков современных систем информирования пассажиров. Ее решение может значительно повысить качество транспортных услуг, оптимизировать работу транспортной системы и улучшить удовлетворенность пассажиров общественным транспортом.

Существующие решения задачи получения данных о загруженности транспорта в режиме реального времени основаны на нескольких технологических подходах [2, 3].

Для оперативного получения точных данных о пассажиропотоке могут применяться различные типы датчиков [4]. Одни из них предназначены для подсчета количества вошедших и вышедших пассажиров – они устанавливаются на входах и выходах транспортного средства. Другие датчики измеряют вес транспортного средства, анализируя тем самым изменение общей массы. Третий тип датчиков отслеживает давление на пол и посадочные места. Несмотря на эффективность в сборе точных данных, применение таких устройств связано со значительными затратами. Среди основных расходов можно назвать стоимость самих датчиков, затраты на их монтаж и интеграцию с существующими системами, а также регулярные расходы на техническое обслуживание и калибровку оборудования.

Интеграция мониторинга загруженности транспорта с системами оплаты проезда дает информацию о количестве проданных билетов и тем самым позволяет оценить часть пассажиропотока. Однако этот метод имеет существенные ограничения. Во-первых, он не учитывает пассажиров, которые проезжают без билета, что может значительно исказить реальную картину загруженности. Во-вторых, в случаях, когда стоимость билета не зависит от дальности поездки, данные о проданных билетах не отражают фактическую загруженность транспорта на разных участках маршрута – пассажиры могут покинуть транспорт до конечной остановки.

Подходы, основанные на анализе данных мобильных приложений, отслеживающих геолокацию пассажиров, или подсчете количества подключений к беспроводным сетям в транспортных средствах [4], а также сервисы, где пользователи самостоятельно отмечают заполненность транспорта, демонстрируют низкую эффективность при небольшом количестве активных участников.

Традиционное прогнозирование загруженности по историческим данным [5–7] затруднено из-за отсутствия точной информации о заполненности транспорта в про-

шлом. Кроме того, на пассажиропоток влияет множество переменных факторов, таких как погода, обстановка на дорогах и другие параметры [6, 8].

Рассматриваемый в данной работе подход на основе интеллектуального анализа видео из салонов транспортных средств лишен перечисленных недостатков: не требует активного участия со стороны пассажиров, эффективен при любой системе оплаты проезда, любом количестве пассажиров, не нуждается в учете дополнительных факторов и часто не требует дополнительных затрат на установку оборудования, так как во многих городах из соображений безопасности транспорт уже оснащен камерами видеонаблюдения [9], данные которых в режиме реального времени передаются в информационные системы организаторов перевозок.

Применение алгоритмов машинного обучения для решения задачи оценки пассажиропотока показывает хорошие практические результаты [10, 11]. При этом могут использоваться классические методы машинного обучения [12], например градиентный бустинг [13, 14], который строит модель последовательно, итеративно добавляя новые модели, чтобы исправить ошибки предыдущих. Также могут применяться методы глубокого обучения [15] и сверточные нейронные сети [16, 17]. В ряде работ встречаются рекуррентные нейронные сети с памятью LSTM [18], содержащие обратные связи и позволяющие сохранять информацию, учитывая долгосрочные зависимости данных, а также гибридные модели [19].

Следует отметить, что в ряде исследований [20, 21] также применяется интеллектуальный анализ видео для оценки загруженности транспорта, при этом акцент делается на подсчете точного количества пассажиров в салоне. Однако для решения задачи определения общей степени загруженности транспортного средства знание абсолютного числа пассажиров не является строго необходимым. Важнее получить обобщенную оценку уровня загруженности, которая позволит оперативно принимать решения по оптимизации транспортных потоков и расписания. Такой подход снижает требования к точности алгоритмов компьютерного зрения, упрощает процесс анализа данных и делает систему более устойчивой [22] к различным внешним факторам (плохая освещенность, затенение, частичное перекрытие обзора). В то же время он сохраняет практическую ценность получаемой информации для информирования пассажиров, планирования расписания и регулирования транспортной нагрузки.

Цель исследования – создание программно-аналитической системы для автоматической оценки степени заполненности автотранспортных средств на основе видеоданных, получаемых с камер, установленных в салонах пассажирского транспорта.

Материал и методы исследования

Входными данными системы являются 8-разрядные монохромные статичные графические файлы разрешением 1280×720, извлеченные из видеопотоков, переданных камерами, установленными в салонах транспортных средств.

Выходными данными системы является заполненность салона по шкале от 1 до 5. Для оценки степени заполненности салона используется пятибалльная шкала, применяемая МБУ «Центральная диспетчерская служба городского пассажирского транспорта» г. Кирова: 1 – «свободно» (свободно более половины мест для сидения, доступно более половины стоячих мест), 2 – «небольшая заполненность» (свободно менее половины мест для сидения, доступно более половины стоячих мест), 3 – «средняя заполненность» (доступно менее половины стоячих мест), 4 – «высокая заполненность» (все сидячие и стоячие места заняты, но есть свободное пространство для перемещения), 5 – «тесно» (все места заняты, стоящие пассажиры располагаются плотно, нет свободного пространства для перемещения).

Критерии оценки эффективности включают точность классификации заполненности салона и время, необходимое для извлечения и анализа одного кадра из видеопотока.

Точность классификации оценивается путем сравнения результатов работы системы с данными экспертной оценки. Для этого всем исходным выборкам видеоданных заранее присваиваются экспертные оценки степени заполненности салона для каждого кадра. В пограничных ситуациях, когда эксперты не могут однозначно определить уровень заполненности по шкале от 1 до 5, выборкам изображений присваивается две экспертные оценки: основная, которая используется для обучения, и дополнительная, которая наряду с основной применяется на этапе валидации и тестирования. Определение загруженности системой считается верным при совпадении результата обработки с основной или дополнительной экспертными оценками.

Система должна функционировать в режиме мягкого реального времени, что подразумевает своевременное предоставление результатов анализа с допустимыми задержками, не критичными для практи-

ческого применения. Время извлечения и анализа кадра будет оцениваться путем измерения интервала между началом процесса обработки кадра и получением финального результата классификации заполненности транспортного средства с учетом предполагаемой вычислительной мощности устройств и нагрузки на систему. Для этого определяется тестовый набор видеофайлов, который включает кадры с различной степенью заполненности салона и при разных условиях съемки. Измерение времени обработки производится встроенными в разрабатываемую программно-аналитическую систему инструментами поочередно для каждого видеофайла. После чего определяется среднее и максимальное время обработки одного кадра в зависимости от заполненности салона транспортного средства.

Будем считать систему эффективной, если средняя вероятность точной классификации изображений составляет не менее 88 % при вероятности грубой ошибки классификации не более 1 % (грубой ошибкой считается расхождение результата работы системы с экспертной оценкой более чем на один уровень по пятибалльной шкале), а среднее время извлечения и анализа кадра не превышает заданного порога в 5 с, а отклонения от среднего значения не приводят к неприемлемым задержкам более 10 с в получении результатов.

Исследование предполагает поэтапное решение задач разработки программно-аналитической системы оценки заполненности салонов пассажирского транспорта. На первом этапе осуществляется сбор и предварительная обработка данных, необходимых для анализа. Затем разрабатывается и обучается нейронная сеть – ключевой элемент создаваемой программно-аналитической системы, которая впоследствии будет использоваться для оценки заполненности салонов транспортных средств. Финальный этап включает интерпретацию результатов работы системы и формирование выводов об эффективности использования машинного обучения для определения уровня загруженности пассажирского транспорта на основе анализа видеоданных из салона.

На первом этапе осуществляется подготовка данных для обучения и тестирования модели. Видеоматериалы (рис. 3), которые использованы для формирования выборок, получены у перевозчиков, обслуживающих городские маршруты г. Кирова: АО «Автотранспортное предприятие» и ООО «Азимут».



Рис. 3. Пример кадров, соответствующих разным уровням загруженности (от 1 до 5)
Примечание: составлен авторами

Полученные видеофайлы разделены на три выборки: тренировочную – для обучения нейронной сети, валидационную – для проверки точности предсказаний во время обучения и тестовую – для проверки работоспособности сети по окончании обучения. Для обеспечения репрезентативности каждой группы данных разделение производится случайным образом.

Каждому кадру видеофайлов присваивается экспертная оценка по следующей методике: видеофайлы просматриваются экспертной группой, вручную отмечаются временные промежутки, в которых уровень заполненности салона по выбранной шкале постоянен, всем кадрам из соответствующих временных промежутков присваивается одинаковая экспертная оценка.

На этапе подготовки данных были учтены потенциальные риски, связанные с недостаточной репрезентативностью выборок, а также возможность возникновения дисбаланса классов [22]. Для минимизации этих рисков был проведен анализ распределения выборок по категориям заполненности салона и применены методы балансировки данных: увеличен объем выборки для ред-

ких классов 4 («высокая заполненность») и 5 («тесно»). Также применялась аугментация данных (случайные повороты на небольшой угол, масштабирование и сдвиги, изменение яркости и контрастности кадров) для предотвращения переобучения и повышения обобщающей способности модели.

Для оптимизации работы системы и снижения нагрузки на оперативную память производится предварительная обработка изображений. Каждый из этапов предварительной обработки направлен на уменьшение объема данных и упрощение их последующей обработки.

Предобработка включает сжатие исходных кадров с разрешением 1280×720 пикселей до 128×72 пикселей, что позволяет существенно сократить объем данных без критической потери качества. Также предобработка включает преобразование уменьшенных изображений в тензоры (двумерные массивы данных), где каждая ячейка отражает значение пикселя в диапазоне от 0 до 255, где 0 соответствует черному цвету, а 255 – белому. На заключительном этапе предобработки производится трансформация двумерной матрицы пикселей

в одномерную матрицу значений пикселей, что упрощает дальнейшую обработку данных алгоритмами машинного обучения.

На втором этапе разрабатывается нейронная сеть для классификации изображений. На входы нейронной сети подается одномерный тензор – упорядоченный набор из 9216 (128×72) значений пикселей, преобразованных из исходного изображения, где каждое значение соответствует яркости пикселя. Архитектура сети включает входной слой, три скрытых слоя и выходной слой, состоящий из пяти нейронов по количеству уровней применяемой шкалы оценки заполненности транспортных средств. Сеть является полносвязной, то есть каждый нейрон предыдущего слоя соединен со всеми нейронами следующего слоя, и выполняет задачу классификации изображений.

Результат работы нейронной сети определяется номером узла с наибольшим значением среди всех выходов. Интерпретация результата работы сети заключается в том, что класс, соответствующий узлу с наибольшим значением, считается наиболее вероятным для данного входного тензора.

Далее проводится обучение нейронной сети. Цель обучения – настроить параметры модели таким образом, чтобы минимизировать ошибку предсказаний на основе тренировочных данных. Обучение сети является обучением «с учителем», так как сеть обучается на готовой выборке с экспертными оценками.

В ходе обучения используется отсев нейронов dropout (0.5) – на каждой итера-

ции обучения случайным образом отключается 50% нейронов для борьбы с переобучением и получением более устойчивой сети. Также процесс обучения включает перемешивание тренировочной выборки на каждой итерации. Это делается для того, чтобы обеспечить случайный порядок представления данных и предотвратить возможные смещения в обучении, связанные с исходным порядком выборки.

В процессе обучения происходит вычисление градиента от результата функции потерь. Функция потерь (или функция ошибки) измеряет разницу между предсказанными моделью значениями и фактическими значениями. Градиент показывает направление и скорость изменения функции потерь относительно весов сети. В данной работе в качестве функции потерь использовалась функция кросс-энтропии, которая измеряет расхождение между предсказанными значениями и реальными данными. Если кросс-энтропия велика, это означает, что разница между двумя распределениями велика, а если кросс-энтропия мала, то распределения похожи друг на друга.

Корректировка весов сети в процессе обучения производилась с помощью функции оптимизации Adam. Этот метод оптимизации адаптивно настраивает скорость обучения для каждого параметра, что позволяет ускорить процесс обучения и улучшить сходимость модели. Корректировка производится в направлении, противоположном градиенту, что способствует минимизации функции потерь.

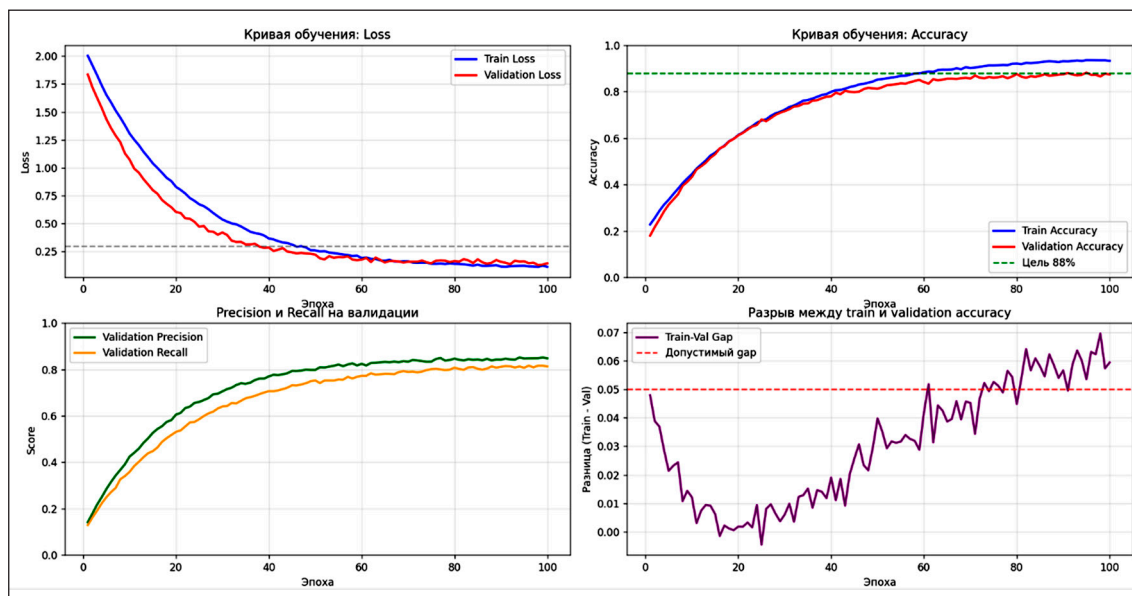


Рис. 4. Графики зависимости метрик качества модели от номера эпохи обучения
Примечание: составлен авторами

После прохождения каждой итерации обучения (эпохи) происходит проверка точности предсказания сети на валидационных данных. При прохождении итераций обучения точность предсказаний растет, но, проходя большое число эпох обучения, точность на валидационных данных падает. Возникает эффект переобучения, то есть нейронная сеть «запоминает» тренировочную выборку данных, соответственно, на данных незнакомых для сети точность становится низкой.

Учитывая риск переобучения [22], необходимо тщательно отслеживать динамику точности предсказаний как на тренировочной, так и на валидационной выборке. Оценка эффективности обучения производится на основе точности предсказаний на валидационной выборке. Для борьбы с переобучением использован метод ранней остановки обучения при достижении плато точности на валидационной выборке. Обучение модели машинного обучения остановлено при достижении значения точности результатов на этапе валидации 88 % (рис. 4).

После определения оптимальных параметров модели проводится проверка обученной сети на тестовой выборке для оценки ее работоспособности и обобщающей способности с использованием программно-аналитической системы.

Результаты исследования и их обсуждение

В рамках исследования разработана программно-аналитическая система, предназначенная для определения уровня загруженности пассажирского транспорта, на основе анализа видеоданных из салонов транспортных средств. Система включает в себя два основных компонента: модуль интеллектуального анализа и пользовательский интерфейс.

Модуль интеллектуального анализа отвечает за обработку видеоданных, полученных из салонов автобусов. В его основе лежит разработанная и обученная многослойная нейронная сеть. Важной особенностью системы является возможность интеграции модуля интеллектуального анализа с существующими информационными системами и базами данных транспортных перевозчиков. Благодаря этому возможна автоматизация процессов получения информации о загруженности транспортных средств, что, в свою очередь, позволяет в реальном масштабе времени информировать пассажиров и оптимизировать работу транспортной системы.

Пользовательский интерфейс разработан как инструмент для исследования эффективности работы обученной нейронной сети. Он позволяет загружать массивы статичных изображений (JPEG, PNG, BMP, TIFF) или видеофайлов (H.264, H.265, MPEG, AVI, MP4), с возможностью добавления поддержки других графических форматов. После загрузки данных пользователь может присвоить или изменить экспертные оценки кадрам или фрагментам видео, а также сравнить результаты машинной классификации с этой оценкой. С помощью встроенных инструментов интерфейс обеспечивает автоматическое определение вероятностей верной классификации и грубой ошибки, а также времени, затраченного на анализ.

Для реализации программного обеспечения были использованы различные инструменты и технологии. В качестве основной среды разработки выбран язык программирования Python, который обладает богатыми возможностями для работы с данными и предлагает широкий спектр библиотек для машинного обучения и анализа данных.

При разработке использовался фреймворк PyTorch. С его помощью было реализовано построение, обучение и тестирование нейронной сети. Фреймворк позволил гибко реализовать архитектуру сети, определить функции потерь и оптимизаторы, а также эффективно работать с тензорами и градиентами.

При предварительной обработке данных, работе с многомерными массивами и матрицами при подготовке их для анализа нейронной сетью применялась библиотека NumPy.

Обработка и анализ видеоданных осуществлялись с помощью библиотеки OpenCV. Библиотека позволила извлечь из видеоданных необходимые признаки и провести предварительную обработку изображений перед подачей на вход нейронной сети.

Основой для разработки пользовательского интерфейса системы стал фреймворк PyQt. С его помощью была создана визуальная оболочка, которая позволяет оператору взаимодействовать с системой, просматривать результаты анализа и управлять процессом работы. Интерфейс разработанной программы представлен на рис. 5.

Оценку эффективности разработанной нейронной сети и программно-аналитической системы для автоматической оценки загруженности пассажирского транспорта будем выполнять по критериям точности recall, precision, f1-score и accuracy.

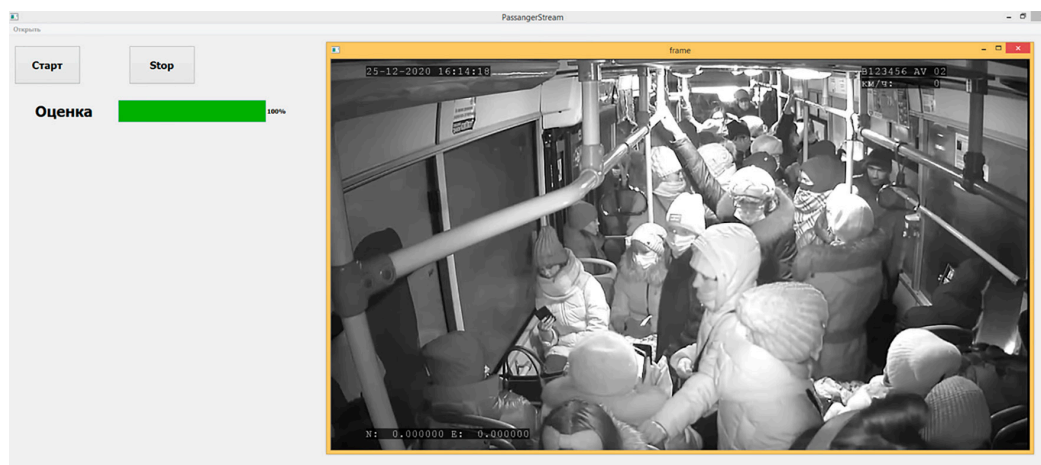


Рис. 5. Интерфейс программы
Примечание: составлен авторами

Recall (полнота) показывает, какая доля реальных случаев определенного уровня загруженности была правильно идентифицирована моделью. Precision (точность определения) показывает, какая доля объектов, выделенных как принадлежащие к определенному классу загруженности, действительно относится к этому классу. F1-score показывает баланс между Precision и Recall (гармоническое среднее точности и полноты). Accuracy (правильность) показывает долю правильно классифицированных объектов от общего числа выборок. При выборе метода усреднения результатов принимается во внимание равнозначность классов и отсутствие дисбаланса классов в тестовой выборке.

Таким образом, для вычисления точности классификации отдельно для уровней загруженности и общего усреднения по классам используются формулы

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i},$$

$$\text{Recall} = \frac{1}{5} \sum_{i=1}^5 \text{Recall}_i,$$

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i},$$

$$\text{Precision} = \frac{1}{5} \sum_{i=1}^5 \text{Precision}_i,$$

$$F1_i = 2 \times \frac{\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i},$$

$$F1 = \frac{1}{5} \sum_{i=1}^5 F1_i,$$

$$\text{Accuracy} = \frac{TP_1 + TP_2 + TP_3 + TP_4 + TP_5}{N},$$

где TP_i (True Positives) – верно предсказанные случаи i -го уровня загруженности, TN_i (True Negatives) – верно предсказанные случаи, не относящиеся к i -му уровню загруженности, FP_i (False Positives) – случаи, ошибочно предсказанные как i -й уровень загруженности, FN_i (False Negatives) – случаи i -го уровня загруженности, которые модель не распознала, N – общее количество тестовых выборок.

Вероятность грубых ошибок PG, определяется по формуле

$$PG_i = \frac{G_i}{N_i},$$

$$PG = \frac{1}{5} \sum_{i=1}^5 PG_i,$$

где G_i – количество расхождений результата работы системы с экспертной оценкой более чем на один уровень для объектов с экспертной оценкой i , N_i – общее количество объектов в выборке с экспертной оценкой i .

Исследование проводилось на тестовой выборке видеофайлов, полученных из различных транспортных средств, с применением разработанной программно-аналитической системы. Для сопоставления оценки эффективности системы при «знакомых» и «незнакомых» нейронной сети условиях были проанализированы две группы тестовых данных.

Таблица 1

Точность классификации видеофайлов из транспорта,
представленного в тренировочной и валидационной выборках

Класс	Recall	Precision	F1-Score	Accuracy	PG
1	92,1 %	89,1 %	90,2 %	–	0,48 %
2	83,7 %	85,2 %	84,3 %	–	0,88 %
3	81,5 %	82,8 %	81,1 %	–	0,86 %
4	83,5 %	84,9 %	83,4 %	–	0,64 %
5	88,7 %	91,5 %	90,1 %	–	0,72 %
Среднее	85,9 %	86,7 %	85,8 %	88,2 %	0,72 %

Примечание: составлена авторами на основе полученных данных в ходе исследования

Таблица 2

Точность классификации видеофайлов
из незнакомых нейронной сети транспортных средств

Класс	Recall	Precision	F1-Score	Accuracy	PG
1	72,6 %	84,9 %	83,1 %	–	11,43 %
2	73,7 %	70,5 %	72,7 %	–	5,89 %
3	73,6 %	72,5 %	73,2 %	–	5,03 %
4	79,3 %	72,8 %	73,8 %	–	4,51 %
5	81,9 %	82,2 %	83,5 %	–	7,04 %
Среднее	76,2 %	76,6 %	76,3 %	76,1 %	6,78 %

Примечание: составлена авторами на основе полученных данных в ходе исследования

В табл. 1 представлена точность классификации для тестовых выборок, снятых видеокамерами в транспортных средствах, характеристики которых (ракурс съемки и планировка салона) были известны модели нейронной сети – видеозаписи из этих типов транспорта присутствовали в тренировочной и валидационной выборках. В табл. 2 отражены вероятности классификации для тестовых выборок, полученных с помощью видеокамер в транспортных средствах другого типа, не представленного в тренировочной и валидационной выборках, с незнакомым нейронной сети ракурсом съемки и планировкой салона.

Средняя вероятность правильной классификации изображений из известных нейронной сети транспортных средств составила 88,2 %, при средней вероятности грубых ошибок 0,72 %. Однако проведенное исследование продемонстрировало значительное снижение точности классификации при использовании модели нейронной сети, обученной на видеоданных, полученных из транспортных средств другого типа. Анализ показал, что ключевые факторы, приво-

дящие к снижению эффективности, заключаются в различиях освещенности, планировке салонов и ракурсе обзора камер. Эти факторы определяют расположение и яркость критически важных для анализа зон нахождения пассажиров на изображении, что, в свою очередь, существенно снижает точность классификации.

Анализ структуры ошибок выявил, что программно-аналитическая система демонстрирует определенные закономерности в некорректной классификации уровней заполненности. Наибольшее количество ошибок (32 % от общего числа ошибок) приходится на случаи, когда система неверно определяет пятый уровень заполненности («тесно»), классифицируя его как четвертый уровень («высокая заполненность»), а также ошибочно идентифицирует четвертый уровень как пятый. 22 % от общего числа ошибок наблюдается, когда система ошибочно идентифицирует второй уровень заполненности («небольшая заполненность») как третий уровень («средняя заполненность») и наоборот – третий уровень ошибочно классифицируется как второй.

Таблица 3

Длительность анализа одного кадра изображения

Класс	Время анализа, мс		
	минимальное	среднее	максимальное
1	19	20,38	22
2	19	20,77	22
3	19	20,04	21
4	20	20,19	21
5	20	20,71	21

Примечание: составлена авторами на основе полученных данных в ходе исследования

Таким образом, основные ошибки классификации касаются смещения соседних уровней заполненности, что указывает на сложности в разграничении близких по характеристикам состояний. Из чего можно сделать вывод, что уменьшение количества классов за счет объединения близлежащих уровней загруженности, позволит существенно увеличить точность классификации.

Основная причина возникновения грубых ошибок классификации – расположение пассажиров близко к камере, что перекрывает часть салона и затрудняет корректный анализ изображения. Подобные случаи демонстрируют, что даже незначительные препятствия в поле зрения камеры могут существенно повлиять на точность работы системы.

Данные времени извлечения и анализа кадра приведены в табл. 3.

Исследование показало, что время анализа не зависит от класса изображения, сложности сцен и иных параметров. Среднее время анализа одного кадра составило 20,43 мс, максимальное – 22, минимальное – 19, что свидетельствует о высокой скорости работы системы и ее пригодности для анализа пассажиропотока в режиме реального времени.

Таким образом, разработанную систему можно считать эффективной для определения уровня загруженности пассажирского транспорта на основе анализа видеоданных из салона в режиме реального времени, но только при условии, что нейронная сеть обучалась на выборках данных, полученных из салонов аналогичных транспортных средств. В случае работы с транспортными средствами иных типов требуется дообучение модели.

Заключение

В ходе исследования была рассмотрена актуальная проблема отсутствия данных

о загруженности общественного транспорта в реальном масштабе времени, что негативно влияет на качество услуг и удовлетворенность пассажиров. Отсутствие такой информации приводит к неэффективному планированию маршрутов и расписания, недогрузке или перегрузке транспорта, потере доходов транспортных компаний, а также затрудняет решение более широких задач развития транспортной системы.

Для решения задачи определения общей степени загруженности транспортного средства был предложен альтернативный подход, основанный на интеллектуальном анализе видео из салонов транспортных средств, который обладает рядом преимуществ по сравнению с существующими технологиями.

В рамках работы была разработана и обучена многослойная нейронная сеть для анализа видеоматериалов и создана прикладная программно-аналитическая система, способная определять уровень загруженности транспортных средств в режиме реального времени.

Количественные результаты исследования демонстрируют высокую эффективность разработанной системы: средняя вероятность правильной классификации изображений составила 88,2%, среднее время анализа одного кадра составило 0,02 с.

Разработанная система может быть успешно применена для решения ряда задач: информирования пассажиров, ожидающих транспорт, о степени его загруженности; повышения качества организации движения городского транспорта путем более эффективного планирования маршрутов и расписания; анализа эффективности пассажирских перевозок и оптимизации транспортной системы в целом.

Таким образом, применение машинного обучения для определения уровня загруженности пассажирского транспорта на основе анализа видеоданных из салона без су-

пешественных материальных затрат позволяет повысить качество транспортных услуг, улучшить удовлетворенность пассажиров и предоставить данные для оптимизации работы транспортной системы.

Список литературы

1. Манаков А.Л., Коларж С.А., Саломатов Е.М. Алгоритм корректировки количества и класса пассажирских транспортных средств на основе данных пассажиропотока // Вестник Сибирского государственного университета путей сообщения. 2024. № 1 (68). С. 13–21. URL: <https://cyberleninka.ru/article/n/algoritm-korrektirovki-kolichestva-i-klassa-passazhirskih-transportnyh-sredstv-na-osnove-dannyh-passazhir-porotoka> (дата обращения: 27.11.2025). DOI: 10.52170/1815-9265_2024_68_13.
2. Silva P., Campos J., Matos J., Salgado J., Salazar C., Costa H., Portela F., Carneiro D. Public Transportation Occupancy Rate // Sustainable, Innovative and Intelligent Societies and Cities. 2023. P. 389–412. URL: https://link.springer.com/chapter/10.1007/978-3-031-30514-6_16 (дата обращения: 27.11.2025). DOI: 10.1007/978-3-031-30514-6_16.
3. Kuchár P. et al. Passenger occupancy estimation in vehicles: A review of current methods and research challenges // Sustainability. 2023. Т. 15. № 2. P. 1332. URL: <https://www.mdpi.com/2071-1050/15/2/1332> (дата обращения: 27.11.2025).
4. Grgurević I., Juršić K., Rajić V. Overview of Wi-Fi-based automatic passenger counting solutions in public urban transport // Sustainable Management of Manufacturing Systems in Industry 4.0. Cham: Springer International Publishing, 2022. P. 181–196. URL: https://link.springer.com/chapter/10.1007/978-3-030-90462-3_12 (дата обращения: 27.11.2025). DOI: 10.1007/978-3-030-90462-3_12.
5. Xue Q. et al. Passenger flow forecasting approaches for urban rail transit: A survey // International Journal of General Systems. 2023. Т. 52. № 8. P. 919–947. URL: <https://www.tandfonline.com/doi/abs/10.1080/03081079.2023.2231133> (дата обращения: 27.11.2025).
6. Saake S., Sommer C. Design of a Forecasting Method for Occupancy Rates in Local Public Transport Based on Data from Automatic Passenger Counting Systems // Interdisciplinary Conference on Production, Logistics and Traffic. Cham: Springer International Publishing, 2023. P. 13–28. URL: https://link.springer.com/chapter/10.1007/978-3-031-28236-2_2 (дата обращения: 27.11.2025).
7. Wood J., Yu Z., Gayah V.V. Development and evaluation of frameworks for real-time bus passenger occupancy prediction // International Journal of Transportation Science and Technology. 2023. Т. 12. № 2. P. 399–413. URL: <https://www.sciencedirect.com/science/article/pii/S2046043022000296> (дата обращения: 27.11.2025).
8. Li P., Wu W., Pei X. A separate modelling approach for short-term bus passenger flow prediction based on behavioural patterns: A hybrid decision tree method // Physica A: Statistical Mechanics and its Applications. 2023. Т. 616. P. 128567. URL: <https://www.sciencedirect.com/science/article/pii/S037843712300122X> (дата обращения: 27.11.2025).
9. Menezes de Á.L.L., Almeida de R.C., Almonfrey D., Nunes R.B. Using Artificial Intelligence for Obtaining Vehicle Occupancy Using Security Cameras // 14th IEEE International Conference on Industry Applications (INDUSCON), São Paulo, Brazil. 2021. P. 635–639. URL: <https://ieeexplore.ieee.org/abstract/document/9529418/> (дата обращения: 27.11.2025).
10. Сеславина Е.А., Гришина А.А. Применение машинного обучения в транспортной отрасли // Цифровая трансформация в экономике транспортного комплекса. М.: ИНФРА-М, 2025. С. 232–236. URL: <https://elibrary.ru/item.asp?id=80617070> (дата обращения: 27.11.2025). DOI: 10.12737/conferencearticle_678931092d40c3.39138565.
11. Zhang J. et al. A real-time passenger flow estimation and prediction method for urban bus transit systems // IEEE Transactions on Intelligent Transportation Systems. 2017. Т. 18. № 11. P. 3168–3178. URL: <https://ieeexplore.ieee.org/abstract/document/7898469/> (дата обращения: 27.11.2025).
12. Park Y. et al. Machine learning approach for study on subway passenger flow // Scientific Reports. 2022. Т. 12. № 1. P. 2754. URL: <https://www.nature.com/articles/s41598-022-06767-7> (дата обращения: 27.11.2025).
13. Lv W. et al. A bus passenger flow prediction model fused with point-of-interest data based on extreme gradient boosting // Applied Sciences. 2022. Т. 12. № 3. P. 940. URL: <https://www.mdpi.com/2076-3417/12/3/940> (дата обращения: 27.11.2025).
14. Zou L. et al. Passenger flow prediction using smart card data from connected bus system based on interpretable XGBoost // Wireless Communications and Mobile Computing. 2022. Т. 2022. № 1. P. 5872225. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/5872225> (дата обращения: 27.11.2025).
15. Nagaraj N. et al. Passenger flow prediction in bus transportation system using deep learning // Multimedia tools and applications. 2022. Т. 81. № 9. P. 12519–12542. URL: <https://link.springer.com/article/10.1007/s11042-022-12306-3> (дата обращения: 27.11.2025).
16. Liu G. et al. Passenger flow estimation based on convolutional neural network in public transportation system // Knowledge-Based Systems. 2017. Т. 123. P. 102–115. URL: <https://www.sciencedirect.com/science/article/pii/S0950705117300849> (дата обращения: 27.11.2025).
17. Lv S. et al. An origin–destination passenger flow prediction system based on convolutional neural network and passenger source-based attention mechanism // Expert Systems with Applications. 2024. Т. 238. P. 121989. URL: <https://www.sciencedirect.com/science/article/pii/S0957417423024910> (дата обращения: 27.11.2025).
18. Balasubramani K., Natarajan U.M. Improving bus passenger flow prediction using Bi-LSTM fusion model and SMO algorithm // Babylonian Journal of Artificial Intelligence. 2024. Т. 2024. P. 73–82. URL: <https://journals.mesopotamian.press/index.php/BJAI/article/view/450> (дата обращения: 27.11.2025).
19. Chen T. et al. Prediction of public bus passenger flow using spatial–temporal hybrid model of deep learning // Journal of Transportation Engineering, Part A: Systems. 2022. Т. 148. № 4. P. 04022007. URL: <https://ascelibrary.org/doi/abs/10.1061/JTEPBS.0000653> (дата обращения: 27.11.2025).
20. Москвичев О., Леонова С., Васильев Д. Использование нейронной сети для учета и анализа показателей работы транспортного предприятия // Мир транспорта. 2023. Т. 21. № 2 (105). С. 39–46. URL: <https://mirtr.elpub.ru/jour/issue/download/75/59#page=41> (дата обращения: 27.11.2025).
21. Zhao J. et al. Detection of passenger flow on and off buses based on video images and YOLO algorithm // Multimedia Tools and Applications. 2022. Т. 81. № 4. P. 4669–4692. URL: <https://link.springer.com/article/10.1007/s11042-021-10747-w> (дата обращения: 27.11.2025).
22. Холл П., Керрис Дж., Пандей П. Машинное обучение для приложений высокого риска: подходы к ответственному искусственному интеллекту. СПб.: БХВ-Петербург, 2025. 320 с. ISBN 978-601-08-4982-2.

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest: The authors declare that there is no conflict of interest.