

УДК 625.033.34

Применение нейронных сетей для распознавания объектов на железнодорожном транспорте

А. Н. Сычугов, В. Н. Михейчиков, М. В. Чернышов

Петербургский государственный университет путей сообщения Императора Александра I, Российская Федерация, 190031, Санкт-Петербург, Московский пр., 9

Для цитирования: *Сычугов А. Н., Михейчиков В. Н., Чернышов М. В.* Применение нейронных сетей для распознавания объектов на железнодорожном транспорте // Известия Петербургского университета путей сообщения. — СПб.: ПГУПС, 2023. — Т. 20. — Вып. 2. — С. 478–491. DOI: 10.20295/1815-588X-2023-2-478-491

Аннотация

Цель: С помощью систем технического зрения и нейросетей, таких как YOLOv8 и MASK R-CNN, можно быстро и точно обнаруживать предметы, которые могут привести к аварии или задержке поездов. YOLOv8 — это один из самых популярных алгоритмов обнаружения объектов в реальном времени, который использует глубокие нейронные сети для классификации и локализации объектов. YOLOv8 может обнаруживать объекты на изображениях и видео с высокой скоростью и точностью. Эта модель может работать на различных аппаратных платформах, включая мобильные устройства и компьютеры. MASK R-CNN — это еще более продвинутый алгоритм обнаружения объектов, который имеет способность выделять объекты и их контуры с высокой точностью. MASK R-CNN использует сверточные нейронные сети и методы сегментации масок для обнаружения объектов. Он может работать как в режиме реального времени, так и на статических изображениях. Когда системы технического зрения оснащены нейросетями YOLOv8 и MASK R-CNN, они могут оперативно реагировать на посторонние объекты, которые появляются на рельсах. Целью статьи является разработка алгоритмов обнаружения объектов железнодорожного транспорта и препятствий с помощью технического зрения и нейронных сетей, оценка эффективности алгоритмов. **Методы:** Алгоритм YOLOv8 основан на архитектуре сверточных нейронных сетей и использует методы обучения с учителем. Эта модель принимает изображение в качестве входных данных и выдает оценки вероятности того, что на изображении присутствует определенный объект в режиме реального времени. Для этого YOLOv8 использует методы определения областей интереса (ROI), позволяющие определить области изображения, на которых могут быть расположены объекты. Алгоритм MASK R-CNN использует более сложные методы, такие как методы сегментации масок и пропорциональное изменение размера области интереса (RoIAlign), что позволяет получать более точные результаты выделения объектов на изображениях и видео. Он также основан на сверточных нейронных сетях и использует методы обучения с учителем. MASK R-CNN использует методы сегментации масок для определения контура объекта на изображении, а также метод RoIAlign, который позволяет получить лучшее качество при обработке различных размеров изображений. Общими математическими методами, которые используются в YOLOv8 и MASK R-CNN, являются методы сверточной нейронной сети, обучения с учителем и оптимизации функции потерь. Они основаны на алгоритмах глубокого обучения, таких как стохастический градиентный спуск и обратное распространение ошибки. **Результаты:** Алгоритм обнаружения посторонних объектов на пути следования подвижного состава с помощью системы технического зрения, расчет оценки качества работы нейронных сетей, сформированы матрицы ошибок, получены результаты обработки нейронных сетей. **Практическая значимость:** Разработан алгоритм обнаружения посторонних объектов на пути следования подвижного состава с помощью системы технического зрения, обучены две нейронные сети для обнаружения объектов железнодорожного транспорта и препятствий на пути.

Ключевые слова: Нейронная сеть, цифровые технологии, подвижной состав, алгоритм, техническое зрение.

1. Анализ развития цифровых технологий на подвижном составе

Цифровые технологии на подвижном составе позволяют повысить эффективность, безопасность и качество эксплуатации подвижного состава с помощью сбора, обработки и анализа данных о его состоянии, местоположении, движении и окружающей среде.

Для реализации цифровой стратегии ОАО «РЖД» создало специализированную дочернюю компанию — ООО «РЖД-Технологии», которая координирует ключевые проекты, разрабатывает цифровые продукты и обеспечивает высокотехнологичные разработки в холдингах [1]. Кроме того, ОАО «РЖД» сотрудничает с ведущими ИТ-компаниями, научными организациями и стартапами в области цифровых инноваций.

Основные направления цифровизации РЖД можно разделить на три группы: цифровая железная дорога, цифровые пассажирские решения и цифровые грузовые решения.

С технологических позиций цифровую железную дорогу можно определить как совокупность цифровых технологий и методов описания инфраструктуры, подвижного состава, перевозочного процесса и базирующихся на них технологий управления движением, обеспечения безопасности и содержания инфраструктуры, которые ориентированы на достижение принципиально новых автоматизированных методов планирования, диспетчерского управления движением, ресурсами и обслуживания пассажиров [2]. К таким технологиям относятся:

Промышленный интернет вещей (Industrial Internet of Things) — это сеть устройств и датчиков, установленных на подвижном составе и инфраструктуре, которые передают данные в реальном времени через беспроводную связь на специализированные платформы для обработки и анализа. IoT позволяет мониторить и контролировать состояние подвижного состава, предотвра-

щать неисправности и аварии, оптимизировать расход топлива и ресурсов, а также управлять логистикой и грузоперевозками.

Big Data — это большие объемы структурированных и неструктурированных данных, которые собираются из разных источников, таких как IoT, видеокамеры, радары, спутниковая навигация и другие. Big Data позволяет проводить комплексный анализ данных о подвижном составе и его окружении, выявлять закономерности и зависимости, прогнозировать поведение и потребности клиентов, а также принимать обоснованные решения на основе данных.

Интеллектуальные управляющие системы — это системы, которые используют алгоритмы искусственного интеллекта, машинного обучения и нейронных сетей для автоматизации и оптимизации процессов управления подвижным составом. Интеллектуальные управляющие системы позволяют реализовывать функции помощи машинисту, автономного вождения, адаптивного торможения, а также обеспечивать взаимодействие подвижного состава с инфраструктурой и другими участниками движения.

2. Опыт применения нейронных сетей и машинного зрения на железнодорожном транспорте

Нейросетевые технологии и технологии машинного зрения являются направлениями искусственного интеллекта, позволяющими создавать системы, способные обучаться на данных и распознавать образы, объекты, лица, речь и текст. Как и во многих других отраслях, нейронные сети и машинное зрение нашли применение на железнодорожном транспорте.

В 2019 году ОАО «РЖД» получило первый скоростной электропоезд «Ласточка», оборудованный системой машинного зрения и комплексом специализированных датчиков для получения информации об окружающей обстановке [3].

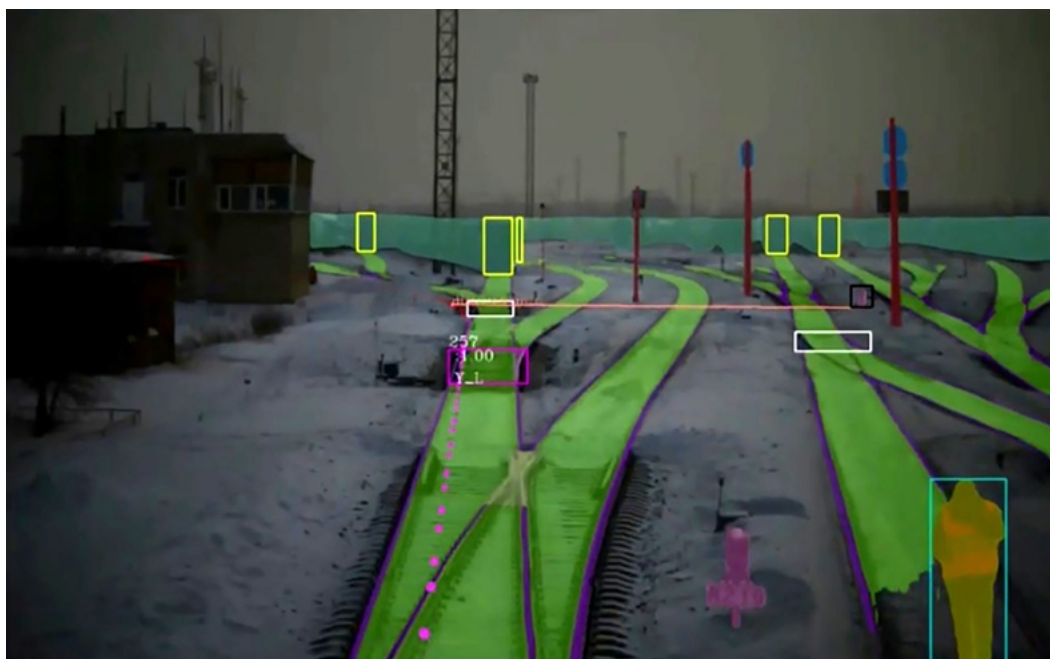


Рис. 1. Обнаружение объектов железнодорожного пути с помощью системы Cognitive Rail Pilot [5]

В состав электропоезда входит комплекс специализированных датчиков, включающий в себя видеокамеры, тепловизоры и лидары. Благодаря этим устройствам электропоезд может получать информацию об окружающей обстановке, что позволяет автоматически контролировать скорость и расстояние до других поездов, а также обеспечивать безопасность пассажиров. Эта технология является одной из инновационных разработок РЖД в области цифровизации транспортной отрасли.

В 2020 году был успешно протестирован прототип сервиса «Цифровой помощник маневрового диспетчера», использующий искусственные нейросети. Этот сервис может собирать и обрабатывать данные из информационных систем станции, формировать варианты роспуска вагонов и выбирать наиболее оптимальный [4]. Результаты тестирования показали, что нейросеть, обученная на примерах работы диспетчеров станции, успешно справляется с роспуском на уровне опытного специалиста. Это означает, что в будущем искусственный интеллект может стать заменой диспетчера в некоторых задачах, освобождая

его от необходимости обрабатывать большое количество информации и повышая общее качество работы. Важно отметить, что искусственный интеллект продолжит развиваться благодаря самообучению.

Компания Cognitive Pilot (дочернее предприятие Сбера и Cognitive Technologies) 25 января 2022 года сообщила о начале серийного производства системы помощи машинисту маневрового железнодорожного локомотива на базе технологий искусственного интеллекта Cognitive Rail Pilot [5]. Благодаря использованию искусственного интеллекта и технического зрения, эти комплексы могут обнаруживать объекты на железнодорожных путях, такие как другие составы, стрелки, пути, люди, светофоры и др. (рис. 1). Это существенно снижает риски возникновения чрезвычайных ситуаций, связанных с человеческим фактором.

В данный момент интеллектуальные комплексы Cognitive Rail Pilot установлены на маневровых локомотивах ЧМЭЗ. Система инициирует автоторможение по светофору, запрещающему положению пошерстной стрелки, по подъезду к

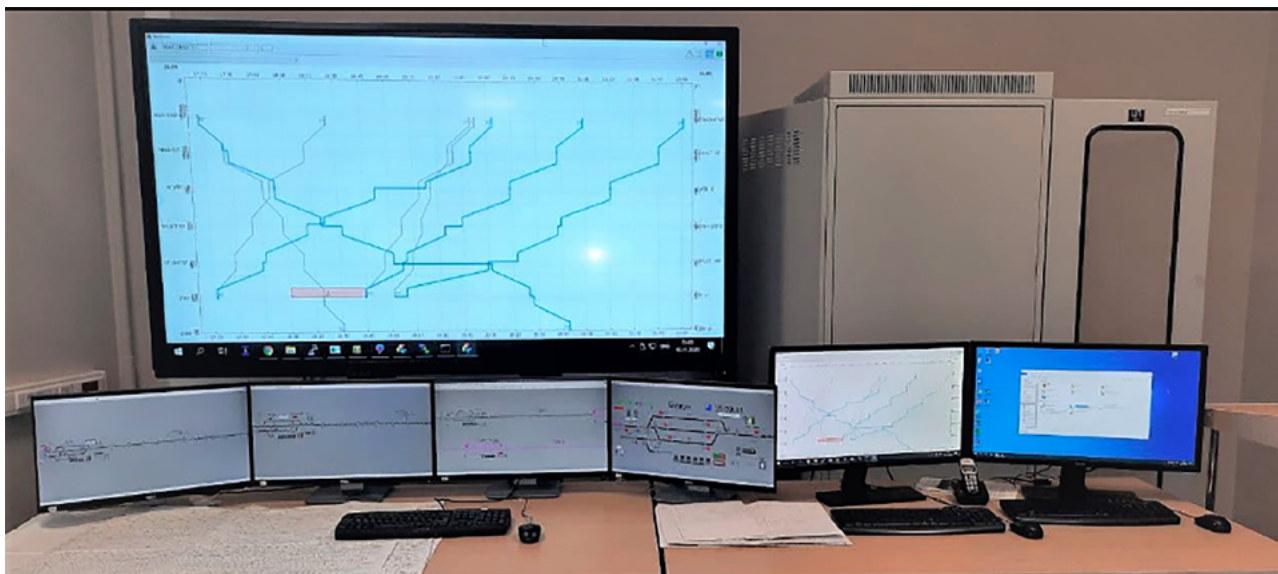


Рис. 2. Рабочее место диспетчера с ИСУДП «Прогноз» [7]

составу быстрее допуска или при препятствии на путях; высчитывает прогноз столкновения, траектории объектов, скорость поезда [3].

Компания планирует завершить дополнительный этап испытаний и установить Cognitive Rail Pilot на пассажирские поезда РЖД к 2024 году. Кроме того, планируется начать экспорт систем в зарубежные страны.

ИСУДП «Прогноз» — это многофункциональная интеллектуальная система прогнозирования, моделирования и оптимизации движения железнодорожного транспорта, в том числе с применением технологий анализа больших данных, разработанная российскими учеными из Центра компетенций НТИ «Искусственный интеллект» на базе МФТИ [6]. Система помогает диспетчерам принимать оперативные и эффективные решения в сложных и конфликтных ситуациях, а также формировать оптимальные и прогнозные графики движения. ИСУДП «Прогноз» повышает безопасность и качество железнодорожных перевозок, сокращает издержки и увеличивает пропускную способность магистралей. Рабочее место диспетчера представлено на рис. 2.

В 2022 году система была установлена в диспетчерском центре г. Алма-Аты на участке Жеты-

ген — Алтынколь. При дальнейшем развитии система сможет в полной мере применяться для построения нормативных графиков движения, а также при интеграции с другими системами планировать оптимальную подвязку локомотивов к поездам, оптимально подбирать время для проведения ремонтных работ с наименьшими потерями для нормативного графика [8].

3. Краткая характеристика основных моделей для классификации и анализа изображений

Задача обнаружения объектов является одной из самых сложных и важных задач машинного зрения, так как она требует одновременно решать задачи классификации изображений (определение типа объекта) и локализации объектов (определение координат объекта).

Для решения задачи обнаружения объектов существует множество моделей нейронных сетей, которые можно разделить на две основные группы: одноуровневые (one-stage) и двухуровневые (two-stage) модели. Одноуровневые модели пытаются решить задачу детекции за один проход по изображению, предсказывая ограничивающие рамки и классы объектов для всех возможных областей

изображения. Двухуровневые модели сначала генерируют кандидаты в ограничивающие рамки (например, с помощью алгоритма региональных предложений), а затем уточняют их положение и класс с помощью отдельной нейронной сети.

Одноуровневые модели имеют преимущества в скорости работы и простоте реализации, но часто уступают двухстадийным моделям в точности детекции, особенно для маленьких или перекрывающихся объектов. Двухуровневые модели имеют преимущества в точности детекции и гибкости настройки, но требуют больше вычислительных ресурсов и времени для работы.

В этом разделе мы кратко охарактеризуем некоторые из самых популярных и эффективных моделей для детекции объектов на изображениях.

R-CNN (Region-based Convolutional Neural Networks) — модель, разработанная в 2014 году компанией Microsoft Research Asia.

Region-based означает использование моделью одноуровневого регионального подхода, то есть вместо обработки всего изображения сразу происходит его разделение на отдельные регионы. Как уже было сказано, такой подход повышает качество распознавания, но значительно понижает быстродействие, так как каждый регион прогоняется через вычислительно сложную сверточную нейронную сеть, что занимает довольно много времени.

В 2015 году той же компанией была выпущена Fast R-CNN, использующая одну и ту же сверточную нейронную сеть для извлечения признаков из всех регионов изображения, а не для каждого региона отдельно, как это делается в R-CNN.

В 2017 году компанией Facebook AI Research была выпущена Mask R-CNN, добавляющая возможность сегментации объектов на изображении, что позволяет получать более точную информацию о расположении объектов.

Все далее рассматриваемые модели являются двухуровневыми.

EfficientNets — семейство моделей нейронных сетей, разработанных компанией Google Brain в 2019 году. Модели достигают высоких результатов на задачах классификации изображений при меньшем количестве параметров, чем другие модели.

Сверточные слои с многомасштабной декомпозицией — это особенность EfficientNets, которая позволяет сети более эффективно использовать информацию изображения на разных масштабах. Для этого сверточные слои в EfficientNets разделяются на несколько подслоев с разными размерами ядра.

Основное преимущество EfficientNets заключается в их высокой эффективности по отношению к вычислительным затратам, что делает их особенно привлекательными для использования на устройствах с ограниченными вычислительными ресурсами, таких как мобильные устройства и встроены системы.

YOLO (You Only Look Once) — модель, разработанная в 2016 году Джозефом Редмоном и Али Ризой Фарди. Название дает понять, что алгоритм является одноуровневым, обрабатывая все изображение за один проход и выполняя классификацию и локализацию объектов одновременно. На момент выпуска модель продемонстрировала отличную скорость обработки изображения в режиме реального времени со скоростью более 45 кадров в секунду на современном оборудовании.

С течением времени были разработаны новые версии YOLO, которые обладают большей скоростью и качеством распознавания. Последней на данный момент вышедшей версией является YOLOv8, выпущенная в начале 2023 года.

VGG (Visual Geometry Group) — семейство моделей, разработанное исследовательской группой Visual Geometry Group в Оксфордском университете в 2014 году.

VGG имеет очень глубокую архитектуру с 16–19 слоями свертки и полносвязанными сло-

ями, что позволяет ей извлекать более высокоуровневые признаки изображения, улучшая точность. При этом все сверточные слои имеют сравнительно малые фильтры размером 3×3 , что позволяет использовать меньшее количество параметров и уменьшить вычислительную сложность модели, тем самым улучшая быстродействие.

VGG часто используется как основная модель для предобучения, которая может быть дообучена на других задачах, например в сетях YOLO для обнаружения объектов.

Существует несколько модификаций архитектуры VGG, которые различаются числом сверточных слоев и их конфигурацией. Например, VGG16, где 16 — количество сверточных слоев.

ResNets — это семейство нейронных сетей, использующих остаточные блоки для решения проблемы затухания градиента при обучении глубоких сетей, предложенные в 2015 году исследователями из Microsoft Research. Авторы также представили различные варианты ResNets с разной глубиной и шириной, такие как ResNet-18, ResNet-50 и другие.

ResNets имеют относительно небольшое количество параметров по сравнению с другими глубокими сетями. Например, ResNet-50 имеет около 25 миллионов параметров, в то время как VGG-16 имеет около 138 миллионов параметров.

Скорость и точность этой модели находятся на достаточно высоком уровне, так как она использует эффективные операции свертки и пакетной нормализации.

4. Алгоритм обнаружения посторонних объектов на пути следования подвижного состава с помощью системы технического зрения

Одной из актуальных задач, решаемых с помощью систем технического зрения на железных дорогах, является обнаружение препятствий и посторонних объектов на рельсах, которые могут

создавать опасность для движения поездов. Эта задача особенно важна при увеличении скоростей движения поездов и роста пассажиропотока, когда машинисту бывает сложно правильно оценить сложившуюся ситуацию и принять оперативное решение.

Первым этапом алгоритма обнаружения является предобработка изображения. Изображения, могут быть зашумленными, размытыми, иметь низкое разрешение или недостаточное освещение, что еще больше усложняет и без того непростую задачу анализа изображения, понижая скорость и точность распознавания. Этот этап включает в себя такие операции, как:

Сбор и хранение изображений с камер. Изображения должны быть достаточно четкими и иметь хорошее освещение. Они должны быть сохранены в формате JPEG или PNG с разрешением не менее 640×480 пикселей.

Нормализация изображений. Для уменьшения влияния различий в освещении, контрасте и цвете изображений, а также для ускорения обработки изображений необходимо привести их к стандартному виду. В процессе нормализации изображением могут быть произведены следующие операции: преобразование изображений в оттенки серого, выравнивание гистограммы яркости, масштабирование изображений до одинакового размера, вычитание среднего значения пикселей по всему набору данных.

Разметка изображений. Для каждого изображения необходимо указать координаты ограничивающих прямоугольников (bounding boxes) вокруг посторонних объектов на рельсах, а также метки классов этих объектов (человек, транспортное средство, железнодорожная колея и т. д.). Разметку можно выполнять вручную с помощью специальных программ или автоматически с помощью алгоритмов компьютерного зрения. Разметка нужна для обучения и тестирования нейронной сети.

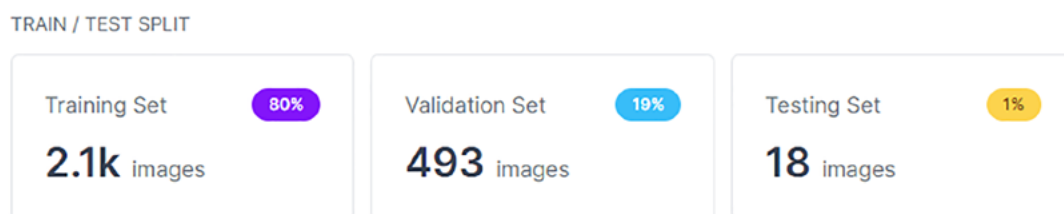


Рис. 3. Количественное разделение набора данных для обучения YOLOv8 [9]

Аугментация изображений. Для увеличения количества и разнообразия данных для обучения нейронной сети, а также для повышения ее устойчивости к различным искажениям изображений можно применять аугментацию изображений. Аугментация заключается в применении случайных преобразований к исходным изображениям, таких как поворот, сдвиг, отражение, изменение цвета, добавление шума и т. д. Аугментация позволяет симулировать различные ситуации, которые могут возникнуть в реальных условиях.

Вторым этапом алгоритма является обучение и тестирование нейронной сети, выполняющей задачу обнаружения посторонних объектов на рельсах. Для этого необходимо выполнить следующие действия:

Выбор архитектуры нейронной сети. Существует множество различных архитектур нейронных сетей, подробнее о которых было рассказано выше. Для демонстрации работы алгоритма и более наглядного сравнения различий будут использоваться две модели с принципиально разными архитектурами: двухуровневая Mask R-CNN и одноуровневая YOLOv8.

Подготовка данных для обучения и тестирования нейронной сети. Для обучения и тестирования нейронной сети необходимо разделить набор данных на две части: обучающую и тестовую. Обучающая часть используется для настройки весов нейронной сети таким образом, чтобы минимизировать функцию потерь, которая измеряет разницу между предсказанными и истинными ограничивающими прямоугольниками и метками классов объектов. Тестовая часть используется

для оценки качества работы нейронной сети на новых данных, которые не были использованы в процессе обучения. Обычно обучающая часть составляет около 80 % от всего набора данных, а тестовая — около 20 %. Для улучшения качества обучения можно также использовать валидационную часть данных, которая отделяется от обучающей части и используется для проверки результатов обучения на каждой эпохе (итерации).

Для обучения модели нейронной сети YOLOv8 нами был подготовлен набор из 2500 фотографий, содержащих классифицируемые объекты, такие как железнодорожная колея, люди, транспорт и т. д. На рис. 3 приведено количественное разделение этого набора данных на обучающую (Training Set), валидационную (Validation Set) и тестовую (Testing Set) части.

Обучение нейронной сети. Обучение нейронной сети заключается в настройке весов сверточных и полносвязных слоев таким образом, чтобы минимизировать функцию потерь на обучающих данных. Для этого используется алгоритм стохастического градиентного спуска, который итеративно обновляет веса в направлении антиградиента функции потерь. Для вычисления градиента функции потерь по весам используется метод обратного распространения ошибки (backpropagation).

Обучение нейронной сети может занимать много времени и требовать больших вычислительных ресурсов, особенно если сеть имеет множество слоев и фильтров. Для ускорения обучения и повышения качества распознавания используются различные техники, такие как:

- Использование предобученной нейронной сети. Вместо того, чтобы обучать нейронную сеть с нуля, можно использовать нейронную сеть, которая уже была обучена на большом наборе данных и дообучить ее на специфическом наборе данных для обнаружения посторонних объектов на рельсах. Это позволяет использовать знания, полученные на общих данных, для решения конкретной задачи. Для этого можно заморозить (не обновлять) веса сверточных слоев предобученной сети и обучать только полносвязные слои.

В связи с большой трудоемкостью самостоятельного обучения модели MASK R-CNN из-за ее математически сложного устройства будет использована предобученная модель, предоставленная разработчиками в открытый доступ на своей официальной странице в интернете.

- Использование пакетной нормализации (batch normalization) — метода, который позволяет повысить производительность и стабилизировать работу искусственных нейронных сетей. Суть данного метода заключается в том, что некоторым слоям нейронной сети на вход подаются данные, предварительно обработанные и имеющие нулевое математическое ожидание и единичную дисперсию [10]. Это улучшает стабильность и скорость обучения, а также уменьшает переобучение.

- Использование метода исключения (dropout) — техники, характеризующей исключение определенного процента случайных нейронов на разных итерациях (эпохах) во время обучения нейронной сети [11]. Это предотвращает переобучение, так как заставляет сеть использовать различные комбинации признаков для классификации объектов.

Тестирование нейронной сети. После того, как нейронная сеть обучена на обучающих данных, необходимо проверить ее работу на тестовых данных, которые не были использованы в процессе обучения. Для этого подается изображение на вход сети и вычисляются выходы всех

слоев. Последний слой сети выдает предсказанные ограничивающие прямоугольники и метки классов объектов на изображении. Формулы для расчета из [12]. Для оценки качества работы нейронной сети используются различные метрики, такие как:

- Доля правильных ответов (accuracy) — это доля правильно классифицированных объектов от общего числа объектов на изображении. Точность вычисляется по формуле (1):

$$\text{accuracy} = \frac{TP}{(TP + FP + FN)}, \quad (1)$$

где TP — истинно положительные (true positive), то есть объекты, которые были правильно обнаружены и классифицированы;

FP — ложно положительные (false positive), то есть объекты, которые были ошибочно обнаружены или классифицированы;

FN — ложно отрицательные (false negative), то есть объекты, которые не были обнаружены или классифицированы.

- Полнота (recall) — это доля правильно обнаруженных объектов от общего числа реальных объектов на изображении. Полнота вычисляется по формуле (2):

$$\text{recall} = \frac{TP}{(TP + FN)}. \quad (2)$$

- Точность (precision) — это доля правильно обнаруженных объектов от общего числа предсказанных объектов на изображении. Точность вычисляется по формуле (3):

$$\text{precision} = \frac{TP}{(TP + FP)}. \quad (3)$$

- F-мера (F-measure) — это гармоническое среднее между точностью и полнотой, которое учитывает баланс между ними. F-мера вычисляется по формуле (4):

$$F - \text{measure} = 2 \cdot \text{precision} \cdot \frac{\text{recall}}{(\text{precision} + \text{recall})}. \quad (4)$$

Для сравнения работы двух моделей на одинаковом наборе тестовых данных нами были подобраны 20 изображений железнодорожной инфраструктуры разной степени сложности. Изображения представляют из себя вид с ракурса съемки камеры технического зрения подвижного состава по ходу его движения и различаются погодными условиями, скоростью движения, качеством фокусировки и т. д. Исходное изображение представлено на рис. 4, обработанные нейронными сетями изображения — на рис. 5 и 6.

Так как модель YOLOv8 была дообучена на пользовательском наборе данных, мы имели возможность добавить еще один класс объектов для распознавания и сегментации, такой как рельсовая колея. Исходная модель MASK R-CNN не имеет этого класса, поэтому в дальнейшем расчете метрик качества работы он не будет учтен.

Расчет оценки качества работы нейронных сетей по формулам (1–4), обозначения по первым буквам (например, accuracy = A):

Accuracy:

$$\text{Mask R-CNN: } A = \frac{65}{(65+15+17)} = 0,68.$$

$$\text{Yolov8: } A = \frac{69}{(69+3+15)} = 0,79.$$

Recall:

$$\text{Mask R-CNN: } R = \frac{65}{(65+17)} = 0,79.$$

$$\text{Yolov8: } R = \frac{69}{(69+15)} = 0,82.$$

Precision:

$$\text{Mask R-CNN: } P = \frac{65}{(65+15)} = 0,81.$$

$$\text{Yolov8: } P = \frac{69}{(69+3)} = 0,95.$$

F-measure:

$$\text{Mask R-CNN: } F = 2 \cdot 0,81 \cdot \frac{0,79}{(0,81+0,79)} = 0,79.$$

$$\text{Yolov8: } F = 2 \cdot 0,95 \cdot \frac{0,82}{(0,95+0,82)} = 0,88.$$

Как видно из расчетов, MASK R-CNN уступает по всем показателям. Главной причиной этого является большое количество False Positive — ложно положительных объектов (в 5 раз больше, чем у конкурента). Это может быть исправлено добавлением специального параметра conf, позволяющего задать отображение рамок классификации по совпадению с определенным процентным значением.

Третий этап алгоритма заключается в оценке результатов работы нейронной сети на тестовых данных и анализе ее сильных и слабых сторон. Для этого можно использовать следующие методы:

Визуализация предсказаний нейронной сети. Для наглядного представления работы нейронной сети можно визуализировать предсказанные ограничивающие прямоугольники и метки классов на тестовых изображениях и сравнить их с истинными. Это позволяет увидеть, какие объекты нейронная сеть обнаруживает правильно, а какие нет, а также какие ошибки она допускает.

Построение кривых обучения и валидации. Для оценки процесса обучения нейронной сети можно построить графики зависимости функции потерь и метрик качества от номера эпохи (итерации) обу-



Рис. 4. Исходное изображение

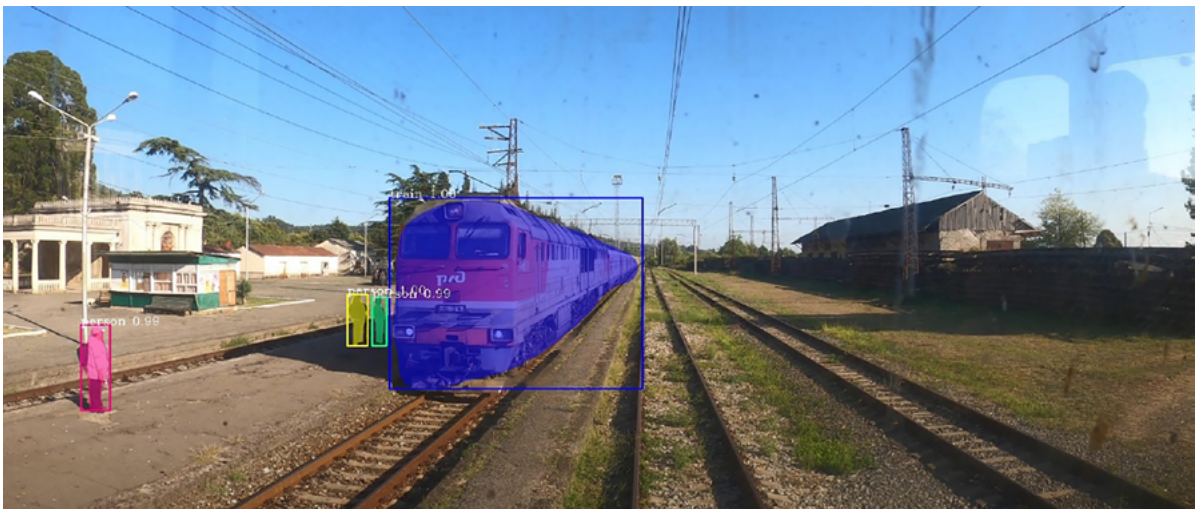


Рис. 5. Изображение, обработанное MASK R-CNN



Рис. 6. Изображение, обработанное YOLOv8

traffic light				82%	20%
car		5%	85%		30%
train		75%			30%
person	64%			4%	20%
background	36%	20%	15%	14%	
	person	train	car	traffic light	background

Рис. 7. Матрица ошибок для Mask R-CNN

traffic light				74%	
car			95%		
train		95%			66%
person	81%				34%
background	19%	5%	5%	26%	
	person	train	car	traffic light	background

Рис. 8. Матрица ошибок для YOLOv8

чения на обучающих и валидационных данных. Это позволяет определить, насколько хорошо нейронная сеть обучается на данных, а также выявить признаки переобучения или недообучения.

Построение матрицы ошибок (confusion matrix). Матрица представляет из себя таблицу, которая позволяет визуализировать эффективность алгоритма классификации путем сравнения прогнозируемого значения целевой переменной с ее фактическим значением. Столбцы матрицы представляют наблюдения в прогнозируемом классе, а строки — наблюдения в фактическом классе (или наоборот) [13]. Например, на рис. 7 и 8 ниже показана матрица ошибок для четырех классов объектов:

Заключение

Результаты, полученные в ходе работы, показывают, что внедрение технического зрения в паре с нейросетевыми технологиями являются эффективным инструментом обнаружения объектов железнодорожного транспорта, людей и препятствий на пути. Предложенные инструменты повышения безопасности в ближайшей перспективе смогут стать одним из элементов комплексной цифровой экосистемы безопасности на железнодорожном транспорте.

Библиографический список

1. РЖД-Технологии. — URL: <https://www.rzdtech.ru/> (дата обращения 21.03.2023).
2. Розенберг Е. Н. Цифровая экономика и цифровая железная дорога / В. И. Уманский, Ю. В. Дзюба // Транспорт Российской Федерации. Журнал о науке, практике, экономике. — 2017. — С. 46. — URL: <https://cyberleninka.ru/article/n/tsifrovaya-ekonomika-i-tsifrovaya-zheleznaaya-doroga/viewer> (дата обращения: 21.03.2023).
3. Машинное зрение на железнодорожном транспорте (РЖД): Cognitive Rail Pilot. — URL: <https://rzdigital.ru/technology/mashinnoe-zrenie/>

(дата обращения: 23.03.2023).

4. Искусственные нейронные сети на железнодорожном транспорте (РЖД): цифровой помощник маневрового диспетчера, Cognitive Rail Pilot. — URL: <https://rzd-digital.ru/technology/iskusstvennye-neyronnye-seti/>.

5. Cognitive Rail Pilot. — URL: https://www.tadviser.ru/index.php/Продукт:Cognitive_Rail_Pilot?ysclid=lfw190xh5o174644139_tadviser.ru.

6. ИСУДП «Прогноз». — URL: <https://pt.2035.university/project/isudp-prognoz>.

7. Система оптимизации движения ж/д транспорта «Прогноз». — URL: https://ai.mipt.ru/projects/sistema_optimizatsii_dvizheniya_zh_d_transporta_prognoz.

8. Расписание за 5 секунд: как нейросеть оптимизирует движение ж/д транспорта. — URL: <https://itnan.ru/post.php?c=1&p=670530>.

9. Roboflow. — URL: <https://app.roboflow.com/>.

10. Batch-normalization. — URL: <https://neerc.ifmo.ru/wiki/index.php?title=Batch-normalization>.

11. Исключение (нейронные сети). — URL: [https://ru.wikipedia.org/wiki/Исключение_\(нейронные_сети\)](https://ru.wikipedia.org/wiki/Исключение_(нейронные_сети)).

12. Метрики в задачах машинного обучения. — URL: <https://habr.com/ru/company/ods/blog/328372/>.

13. Матрица ошибок. — URL: https://help.sap.com/docs/SAP_PREDICTIVE_ANALYTICS/41d1a6d4e7574e32b815f1cc87c00f42/9c144a376f004058b4e9fe56727359af.html?version=3.2.

Дата поступления: 30.05.2023

Решение о публикации: 16.05.2023

Контактная информация:

СЫЧУГОВ Антон Николаевич — ст. преподаватель;
sychugov@pgups.ru

МИХЕЙЧИКОВ Вадим Николаевич — обучающийся;
mihei4ikov@mail.ru

ЧЕРНЫШОВ Максим Витальевич — обучающийся;
mxchrnshv@yandex.ru

Application of Neural Networks for Object Recognition in Railway Transportation

A. N. Sychugov, V. N. Mikheyshikov, M. V. Chernyshov

Emperor Alexander I St. Petersburg State Transport University, 9, Moskovsky pr., Saint Petersburg, 190031, Russian Federation

For citation: Sychugov A. N., Mikheyshikov V. N., Chernyshov M. V. Application of Neural Networks for Object Recognition in Railway Transportation // *Proceedings of Petersburg Transport University*, 2023, vol. 20, iss. 2, pp. 478–491. (In Russian). DOI: 10.20295/1815-588X-2023-2-478-491

Summary

Purpose: With the help of vision systems and neural networks, such as YOLOv8 and MASK R-CNN, it is possible to quickly and accurately detect objects that can lead to an accident or delay trains. YOLOv8 is one of the most popular real-time object detection algorithms that uses deep neural networks to classify and localize objects. YOLOv8 can detect objects in images and videos with high speed and accuracy. This model can work on various hardware platforms, including mobile devices and computers. MASK R-CNN is an even more advanced object detection algorithm that has the ability to highlight objects and their contours with high accuracy. MASK R-CNN uses convolutional neural networks and mask segmentation techniques to detect objects. It can work both in real time and on static images. When vision systems are equipped with YOLOv8 and MASK R-CNN neural networks, they can quickly respond to extraneous objects that appear on the rails. The purpose of the article is to develop algorithms for detecting railway transport objects and obstacles using technical vision and neural networks, and to evaluate the effectiveness of algorithms. **Methods:** The YOLOv8 algorithm is based on the architecture of convolutional neural networks and uses supervised learning methods. This model takes an image as input and provides estimates of the probability that a certain object is present in the image in real time. To achieve this, YOLOv8 employs region of interest (ROI) detection methods, allowing to determine the areas of the image on which objects may be located. The MASK R-CNN algorithm uses more sophisticated methods, such as mask segmentation methods and proportional resizing of the area of interest (RoIAlign) to achieve more accurate results of object detecting in images and videos. It is also based on convolutional neural networks and uses supervised learning methods. MASK R-CNN uses mask segmentation methods to determine the contour of an object in an image, as well as the RoIAlign method, which allows for superior quality when processing various image sizes. Common mathematical methods that are used in YOLOv8 and MASK R-CNN are methods of convolutional neural network, supervised learning and optimization of the loss function. They are based on deep learning algorithms such as stochastic gradient descent and backward propagation of errors. **Results:** An algorithm for detecting foreign objects on the route of rolling stock using a technical vision system, calculation of the evaluation of the quality of neural networks performance, error matrices have been formed, the results of neural network processing have been obtained. **Practical significance:** An algorithm for detecting foreign objects on the route of the moving rolling stock using a technical vision system has been developed, two neural networks have been trained to detect railway transport objects and obstacles on the way.

Keywords: Neural network, digital technologies, rolling stock, algorithm, technical vision.

References

1. *RZhD-Tekhnologii* [RZD-Technologies]. Available at: <https://www.rzdtech.ru/> (accessed: March 21, 2023). (In Russian)

2. Rozenberg E. N., Dzyuba Yu. V. Tsifrovaya ekonomika i tsifrovaya zheleznaya doroga [Digital economy and digital railway]. *Transport Rossiyskoy Federatsii. Zhurnal o nauke, praktike, ekonomike* [Transport of the Russian Federation.

Journal of science, practice, economics]. 2017, p. 46. Available at: <https://cyberleninka.ru/article/n/tsifrovaya-ekonomika-i-tsifrovaya-zheleznaya-doroga/viewer> (accessed: March 21, 2023). (In Russian)

3. *Mashinnoe zrenie na zheleznodorozhnom transporte (RZhD): Cognitive Rail Pilot* [Machine vision in railway transport (RZD): Cognitive Rail Pilot]. Available at: <https://rzdigital.ru/technology/mashinnoe-zrenie/> (accessed: March 23, 2023). (In Russian)

4. *Iskusstvennye neyronnye seti na zheleznodorozhnom transporte (RZhD): tsifrovoy pomoshchnik manevrovogo dispetchera, Cognitive Rail Pilot* [Artificial neural networks in railway transport (RZD): digital assistant to the shunting dispatcher, Cognitive Rail Pilot]. Available at: <https://rzdigital.ru/technology/iskusstvennye-neyronnye-seti/>. (In Russian)

5. Cognitive Rail Pilot. Available at: https://www.tadviser.ru/index.php/Produkt:Cognitive_Rail_Pilot?ysclid=lfw190xh5o174644139_tadviser.ru.

6. *ISUDP “Prognoz”* [ISUDP “Forecast”]. Available at: <https://pt.2035.university/project/isudp-prognoz>. (In Russian)

7. *Sistema optimizatsii dvizheniya zh/d transporta “Prognoz”* [The system for optimizing the movement of railway transport “Prognoz”]. Available at: https://ai.mipt.ru/projects/sistema_optimizatsii_dvizheniya_zh_d_transporta_prognoz. (In Russian)

8. *Raspisanie za 5 sekund: kak neyroset' optimiziruet dvizhenie zh/d transporta* [Schedule in 5 seconds: how a

neural network optimizes the movement of railway transport]. Available at: <https://itnan.ru/post.php?c=1&p=670530>. (In Russian)

9. Roboflow. Available at: <https://app.roboflow.com/>.

10. Batch-normalization. Available at: <https://neerc.ifmo.ru/wiki/index.php?title=Batch-normalization>.

11. *Isklyuchenie (neyronnye seti)* [Exception (neural networks)]. Available at: [https://ru.wikipedia.org/wiki/Isklyuchenie_\(neyronnye_seti\)](https://ru.wikipedia.org/wiki/Isklyuchenie_(neyronnye_seti)). (In Russian)

12. *Metriki v zadachakh mashinnogo obucheniya* [Metrics in machine learning problems]. Available at: <https://habr.com/ru/company/ods/blog/328372/>. (In Russian)

13. *Matritsa oshibok* [Matrix of errors]. Available at: https://help.sap.com/docs/SAP_PREDICTIVE_ANALYTIC_S/41d1a6d4e7574e32b815f1cc87c00f42/9c144a376f004058b4e9fe56727359af.html?version=3.2. (In Russian)

Received: May 16, 2023

Accepted: May 30, 2023

Author’s information:

Anton N. SYCHUGOV — Senior Lecturer;
sychugov@pgups.ru

Vadim N. MIKHEYCHIKOV — Student;
mihei4ikov@mail.ru

Maxim V. CHERNYSHOV — Student;
mxchrnshv@yandex.ru