

СЕГМЕНТАЦИЯ ОБЪЕКТОВ ТЕХНИЧЕСКИМ ЗРЕНИЕМ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ УПРАВЛЕНИЯ ЖЕЛЕЗНОДОРОЖНЫМ ПОДВИЖНЫМ СОСТАВОМ

ФЕДОРОВ Владимир Анатольевич, аспирант; e-mail: Fedorov.Vladimir@UrFU.Ru
ОГОРОДНИКОВА Ольга Михайловна, д-р техн. наук, доцент, профессор кафедры;
 e-mail: Olga.Ogorodnikova@UrFU.Ru

Уральский федеральный университет, кафедра «Электронное машиностроение», Екатеринбург

В статье представлено комплексное исследование конфигураций модели сверточной нейронной сети YOLOv11 применительно к задачам сегментации объектов технического зрением в автоматизированных системах управления железнодорожным подвижным составом. Основное внимание уделено оценке производительности и качества обработки изображений, выполняемой различными конфигурациями модели. Выполнен сравнительный анализ 25 конфигураций YOLOv11, которые различаются архитектурой и разрешением входного изображения в интервале от 640 × 640 до 1920 × 1920 пикселей. Для обучения моделей был создан специализированный набор визуальных данных, в котором 20 000 аннотированных изображений железнодорожной инфраструктуры распределены по 40 классам объектов. Все обученные модели были оценены по точности сегментации с использованием метрики mAP (0,5–0,95) и по производительности. Полученные результаты позволяют выбрать конфигурацию модели YOLOv11 с наиболее подходящими параметрами для использования в составе бортовых систем технического зрения для автоматизированных систем управления железнодорожным подвижным составом в зависимости от предъявляемых требований к точности сегментации объектов и мощности доступных вычислительных ресурсов.

Ключевые слова: техническое зрение; автоматизированная система управления; железнодорожный подвижной состав; обнаружение объектов; сегментация объектов; сверточная нейронная сеть; YOLOv11.

DOI: 10.20295/2412-9186-2025-11-03-239-249

▼ Введение

Современные тенденции в области автоматизации железнодорожного транспорта связаны с развитием интеллектуального аппаратно-программного обеспечения и направлены на достижение высоких уровней автоматизации систем управления подвижным составом [1]. В международной градации, которая поддерживается российским стандартом ГОСТ 72010—2025, высокий третий уровень автоматизации УАЗ предполагает минимальное участие локомотивной бригады и многофункциональность бортового оборудования [2]. При этом система управления должна регулировать интервалы движения составов, контролировать скорость, собирать информацию о приводах, обеспечивать остановку в случае запрещающего сигнала светофора или возникновения препятствий. Наивысший четвертый уровень автоматизации УА4 соответствует полной автоматизации

процессов и предполагает отсутствие локомотивной бригады в поезде [3].

Ожидаемым преимуществом автоматизированных систем управления железнодорожным подвижным составом (АСУ ЖД ПС) высоких уровней является значительный прирост объемов перевозки грузов за счет повышения пропускной способности железнодорожных линий [4] при соблюдении условий безопасности движения за счет исключения человеческого фактора и ошибочных действий машиниста [5].

Построение АСУ ЖД ПС высоких уровней неразрывно связано с использованием систем технического зрения, способных обеспечивать визуальный анализ окружающей обстановки [6]. Задачи визуального анализа включают обнаружение объектов железнодорожной инфраструктуры и выявление потенциально опасных препятствий на пути следования состава. Для обнаружения объекты на изображениях необходимо сегментировать (выделить

и классифицировать) с помощью программного обеспечения. Требования к соответствующему программному обеспечению системы управления регламентируются стандартом ГОСТ 34009—2016 и акцентируют выполнение функций управления единицей тягового подвижного состава в режиме реального времени, в том числе выполнение критически важных сенсорных функций.

Сенсорные функции технического зрения важны при выполнении таких задач автоматического управления, как остановка ЖД ПС в заданной локации маршрута и точное позиционирование относительно пассажирской платформы или предельных столбиков. В разработке находятся также алгоритмы автоматического определения местоположения ЖД ПС в случае отключения спутниковой навигационной системы, например, при прохождении протяженного тоннеля [7].

Цель данной работы — выполнить анализ программного обеспечения с применением модели сверточных нейронных сетей YOLO для системы технического зрения АСУ ЖД ПС и оценить возможность системы в зависимости от конфигурации, обеспечить надежную сегментацию объектов в условиях ограниченных вычислительных ресурсов.

1. Общая характеристика проблем технического зрения АСУ ЖД ПС

Эксплуатация железнодорожного транспорта предполагает круглосуточное движение составов, что формирует специфические проблемы для технического зрения в географических условиях Российской Федерации, в частности, широкий интервал освещенности окружающих объектов и большое разнообразие погодных вариаций.

С учетом высокой изменчивости параметров внешней среды в АСУ ЖД ПС стремятся применять методы обнаружения объектов, основанные на глубоких нейронных сетях и сверточной архитектуре. Так, с использованием сверточных нейронных сетей производилось обнаружение дефектов рельс [8, 9], с помощью модели Faster R-CNN осуществлялась инспекция крепежных элементов железнодорожного полотна [10], модель YOLOv3 (You Only Look

Once) применялась для выявления дефектов контактной сети [11], а с использованием модели YOLOv8 выполнялось обнаружение различных объектов железнодорожной инфраструктуры [12], а также их сегментация [13]. Перечисленные примеры применения нейронных сетей показали повышенную эффективность в сравнении с традиционными способами цифровой обработки изображений [14].

2. Формирование набора данных для обучения сверточных нейронных сетей

Важным условием успешного использования сверточных нейронных сетей является их обучение и формирование наборов данных, содержащих аннотированные изображения объектов. В открытом доступе имеется несколько наборов данных с объектами железнодорожной инфраструктуры: RailSem19, FRSign, RAWPED, Rail-DB, RailSet, GERALD, OSDaR23, RailDataSets. Основная информация о перечисленных наборах данных представлена в табл. 1.

Анализ структуры данных показал, что существующие базы данных не полностью соответствуют отечественным требованиям АСУ ЖД ПС, поскольку визуально отражают зарубежную инфраструктуру. Набор данных RailDataSets, хотя и включает изображения с российских железных дорог, обладает ограниченным разнообразием, поскольку собран из изображений небольшого участка железнодорожного пути.

В данной работе для обучения сверточных нейронных сетей и надежной сегментации объектов железнодорожной инфраструктуры был сформирован собственный набор данных, содержащий 20 000 изображений с разрешением 1920 × 1080 пикселей. Помимо изображений объектов набор данных включает аннотированные данные о местоположении объекта на изображении. Аннотированные данные включают класс объекта, маску сегментации, координаты ограничивающей рамки. В процессе обучения сверточная нейронная сеть обрабатывает входное изображение и сопоставленные с ним аннотированные данные.

Исходные изображения были получены от видеокамеры, установленной на локомотиве, с

Таблица 1. Открытые наборы данных с объектами железнодорожной инфраструктуры

Набор данных	Год выпуска	Количество изображений	Объекты аннотации
RailSem19	2019	8500	Железнодорожная и трамвайная инфраструктура разных стран
FRSign	2020	105 352	Светофоры железной дороги Франции
RAWPED	2020	26 000	Пешеходы на железной дороге разных стран
Rail-DB	2022	7432	Железнодорожные пути на железной дороге Китая
RailSet	2022	7700	Железнодорожные пути на железной дороге Франции
GERALD	2023	5000	Сигналы светофоров на железной дороге Германии
OSDaR23	2023	204 091	Железнодорожная инфраструктура Германии
RailDataSets	2024	8203	Рельсы, сигналы светофоров, препятствия на железной дороге России

разных регионов железной дороги России при различных погодных условиях и в разное время суток. Все изображения были аннотированы для решения задачи сегментации. Всего было размечено 40 классов объектов железнодорожной инфраструктуры (путевые и сигнальные знаки, элементы железнодорожного пути), а также объектов, которые потенциально могут стать динамическими препятствиями (люди, животные, железнодорожный и автомобильный транспорт).

3. Архитектура модели YOLOv11

В данной работе для решения задач обнаружения и сегментации объектов железнодорожной инфраструктуры в бортовых АСУ ЖД ПС использована современная модель сверточной нейронной сети YOLOv11. Анализ архитектуры модели YOLOv11 и ее существующих версий необходим для составления в дальнейшем собственной модификации, ориентированной на специфику обнаружения объектов железнодорожной инфраструктуры бортовой системой управления с ограниченными вычислительными ресурсами.

Архитектура модели YOLOv11 состоит из трех основных частей: скелета, шеи и головы, которые показаны на рис. 1. Скелет (backbone) отвечает за извлечение признаков в трех масштабах из входного изображения (input image). Шея (neck) является промежуточным звеном между скелетом и головой, выполняя агрегацию и улучшение представления извлеченных признаков во всех трех масштабах. Голова (head) выполняет прогнозирование.

Архитектура YOLOv11 использует три специализированные головы (head) для прогнозирования, каждая из которых соответствует своему масштабу признаков: мелкому, среднему и крупному. Каждая голова независимо выполняет предсказание объектов для своего масштаба. Мелкомасштабная голова работает с картами признаков высокого разрешения и лучше детектирует мелкие объекты. Крупномасштабная голова обрабатывает низкоразмерные карты признаков с высокой семантической ценностью и отвечает за обнаружение крупных объектов. Средняя голова обеспечивает компромисс между крайними масштабами. На этапе постобработки предсказания от всех трех голов объединяются и фильтруются для устранения дублирующих детекций, после чего формируется единый финальный набор предсказаний для анализируемого изображения. Финальный отчет содержит выходные данные о множестве обнаруженных объектов, включая информацию о классе и уверенности, о маске объектов и координатах ограничивающих рамок.

Модель YOLOv11 по масштабированию разделяется на пять версий: **nano**, **small**, **medium**, **large** и **extra-large**. Масштабирование осуществляется по трем параметрам: по глубине сети d , по ширине сети w и по максимальному количеству каналов в блоке mc . Параметр глубины d нейронной сети отвечает за количество слоев или повторений модулей; соответственно, чем выше данный параметр, тем глубже сеть, то есть больше задействовано уровней абстракции признаков. Параметр ширины w нейронной сети задает количество каналов в

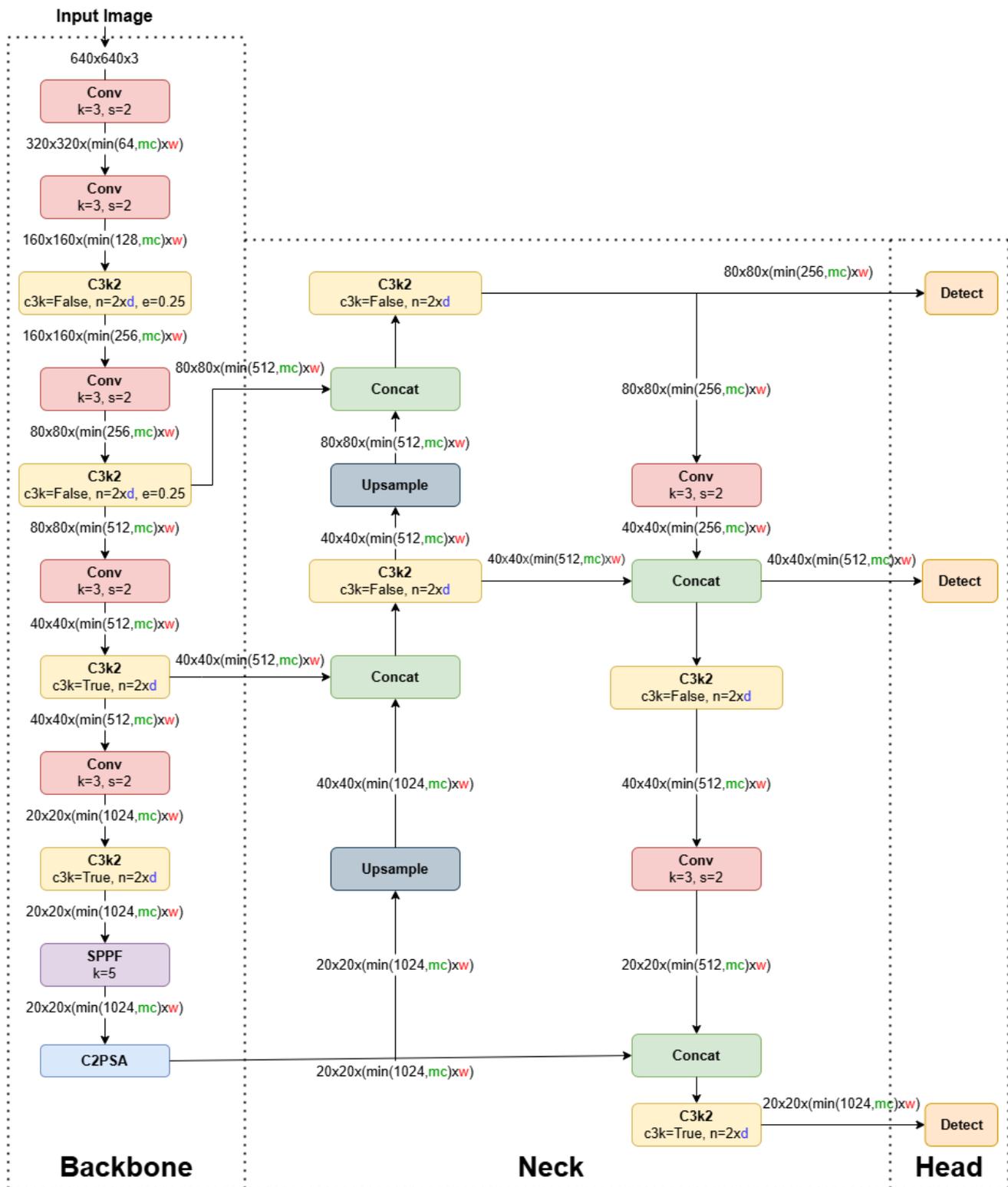


Рис. 1. Архитектура модели YOLOv11

свертках, поэтому чем выше данный параметр, тем больше признаков анализируется на каждом уровне. Параметр максимального количества каналов mc ограничивает число выходных каналов в слоях. В табл. 2 представлены

основные параметры масштабирования для разных версий YOLOv11.

Версии n и s модели YOLOv11 характеризуются параметрами масштабирования, предопределяющими низкую вычислительную

Таблица 2. Параметры масштабирования различных версий модели YOLOv11

Версия модели	<i>d</i> (depth multiple)	<i>w</i> (width multiple)	<i>mc</i> (max channels)
n (nano)	0,50	0,25	1024
s (small)	0,50	0,50	1024
m (medium)	0,50	1,00	512
l (large)	1,00	1,00	512
x (extra-large)	1,00	1,50	512

сложность и малую точность сегментации. Версии *m*, *l* и *x* требуют значительного наращивания вычислительных ресурсов, что ставит под сомнение возможность их развертывания на процессорах бортовой системы управления.

4. Конфигурации YOLOv11 и их обучение

С целью оценки эффективности YOLOv11 в решении задач сегментации объектов по требованиям АСУ ЖД ПС было выполнено обучение 25 конфигураций модели. Сравнению подвергнуты конфигурации, которые отличаются версией модели (**n**, **s**, **m**, **l**, **x**) и разрешением входного слоя (640 × 640, 960 × 960, 1280 × 1280, 1600 × 1600, 1920 × 1920).

Обучение моделей производилось на графическом процессоре NVIDIA RTX A5000 с использованием сформированного в данной работе набора данных. Исходный набор данных перед обучением был разделен на три неравные выборки — тренировочную, валидационную и тестовую — в соотношении 70 %, 20 и 10 % соответственно.

Тренировочная выборка содержит основную часть изображений, которые были использованы непосредственно для обновления весовых коэффициентов сверточной нейронной сети во время обучения с минимизацией функции потерь и настройкой внутренних параметров модели на распознавание характерных признаков объектов.

Валидационная выборка использовалась также на этапе обучения для контроля за обобщающей способностью модели. Следует уточнить, что на валидационной выборке обучение не производилось, данная выборка использовалась исключительно для оценки метрик качества обучения после каждой эпохи.

Тестовая выборка составлена из относительно небольшого количества изображений

и предназначена для окончательной оценки качества модели после завершения обучения. Следует подчеркнуть, что данные из тестовой выборки не были использованы ни в процессе обучения моделей, ни при их валидации, что позволяет получить независимую оценку эффективности различных конфигураций модели YOLOv11 при обработке новых изображений.

Обучение сверточной нейронной сети выполняется итерационно и проходит через следующие шаги:

- прямое распространение входных данных и вычисление предсказаний с использованием текущих параметров модели;
- вычисление суммарной функции потерь, которая учитывает отклонение предсказанных значений от истинных;
- обратное распространение ошибки и вычисление градиентов функции потерь по параметрам модели;
- обновление параметров модели с использованием метода стохастического градиентного спуска.

Все перечисленные шаги циклически повторяются на протяжении заданного числа эпох. Для каждой из конфигураций модели YOLOv11 в процессе обучения устанавливалось максимальное количество эпох, равное 400. С целью снижения вычислительной емкости в алгоритм введен ограничивающий фактор, благодаря которому фактическое количество эпох до завершения обучения оказывалось меньше заданного максимального значения. Эффект экономии расчетного времени связан с использованием механизма ранней остановки, который позволяет предотвратить переобучение модели и сократить объем вычислений при обучении модели. Механизм ранней остановки основывается на мониторинге

Таблица 3. Точность и время сегментации различных конфигураций модели YOLOv11

Версия модели	Разрешение изображений, пиксели	Время сегментации, мс	Метрика качества сегментации mAP (0,5–0,95)
YOLOv11n (nano)	640 × 640	8,2	0,429
	960 × 960	14,5	0,562
	1280 × 1280	24,6	0,644
	1600 × 1600	36,5	0,682
	1920 × 1920	55,2	0,707
YOLOv11s (small)	640 × 640	10,7	0,480
	960 × 960	20,9	0,592
	1280 × 1280	35,3	0,667
	1600 × 1600	52,6	0,712
	1920 × 1920	80,9	0,729
YOLOv11m (medium)	640 × 640	16,3	0,530
	960 × 960	32,1	0,615
	1280 × 1280	54,8	0,692
	1600 × 1600	84,0	0,731
	1920 × 1920	125,0	0,747
YOLOv11l (large)	640 × 640	18,2	0,548
	960 × 960	38,9	0,636
	1280 × 1280	62,1	0,718
	1600 × 1600	94,6	0,756
	1920 × 1920	137,4	0,772
YOLOv11x (extra large)	640 × 640	27,0	0,560
	960 × 960	53,2	0,650
	1280 × 1280	90,9	0,743
	1600 × 1600	144,4	0,778
	1920 × 1920	213,4	0,795

метрики качества сегментации mAP (0,5–0,95) [15], который выполняется на валидационном наборе данных после каждой эпохи. Если в процессе обучения не наблюдается значительное улучшение выбранной метрики, процесс обучения автоматически прерывается. Такой подход позволяет зафиксировать параметры обученной модели в тот момент, когда возникает риск переобучения на тренировочном наборе данных.

После обучения была произведена оценка качества сегментации на тестовой выборке с использованием метрики mAP (0,5–0,95), а также оценка производительности по параметру времени сегментации одного кадра на графическом процессоре NVIDIA RTX A5000. Оценочные параметры приведены в табл. 3.

На рис. 2 представлен график зависимости точности сегментации от разрешения входного слоя различных версий модели YOLOv11. Представленный график демонстрирует рост точности сегментации с увеличением разрешения выходного слоя и с усложнением версии модели. Наибольшую точность сегментации $mAP(0,5-0,95) = 0,795$ продемонстрировала конфигурация модели версии YOLOv11x с разрешением входного слоя 1920 × 1920 пикселей.

На рис. 3 представлен график зависимости времени сегментации одного кадра на графическом процессоре NVIDIA RTX A5000 от разрешения входного слоя разных версий модели YOLOv11. Представленный график демонстрирует повышение вычислительной сложности с увеличением разрешения

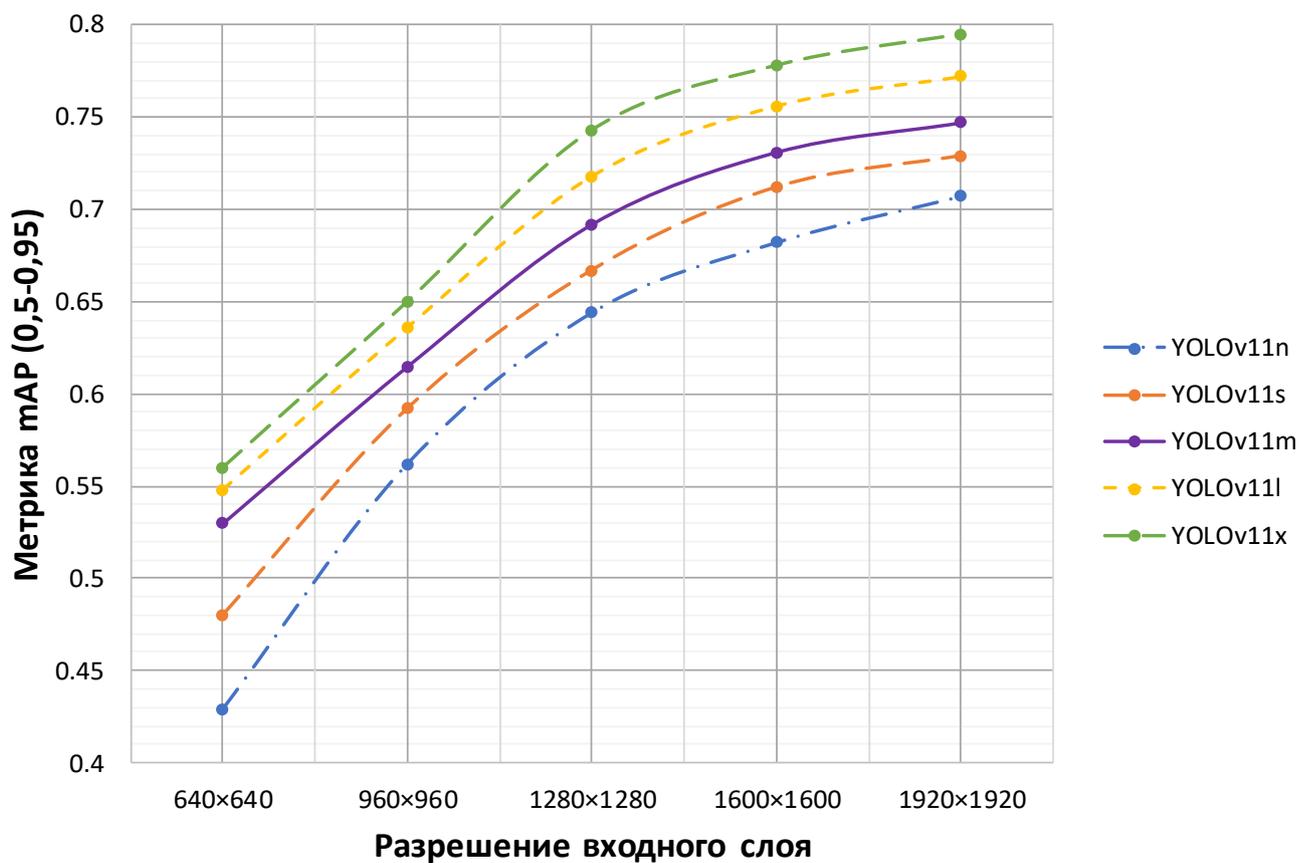


Рис. 2. График зависимости точности сегментации от разрешения

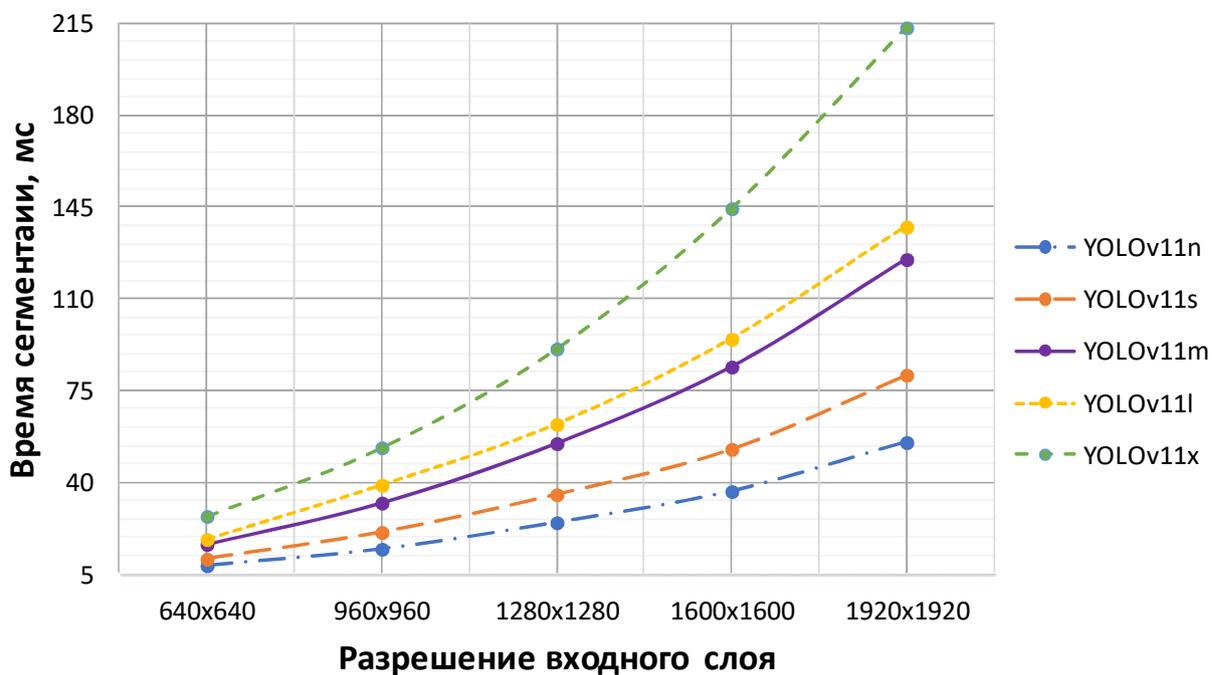


Рис. 3. График зависимости времени сегментации от разрешения

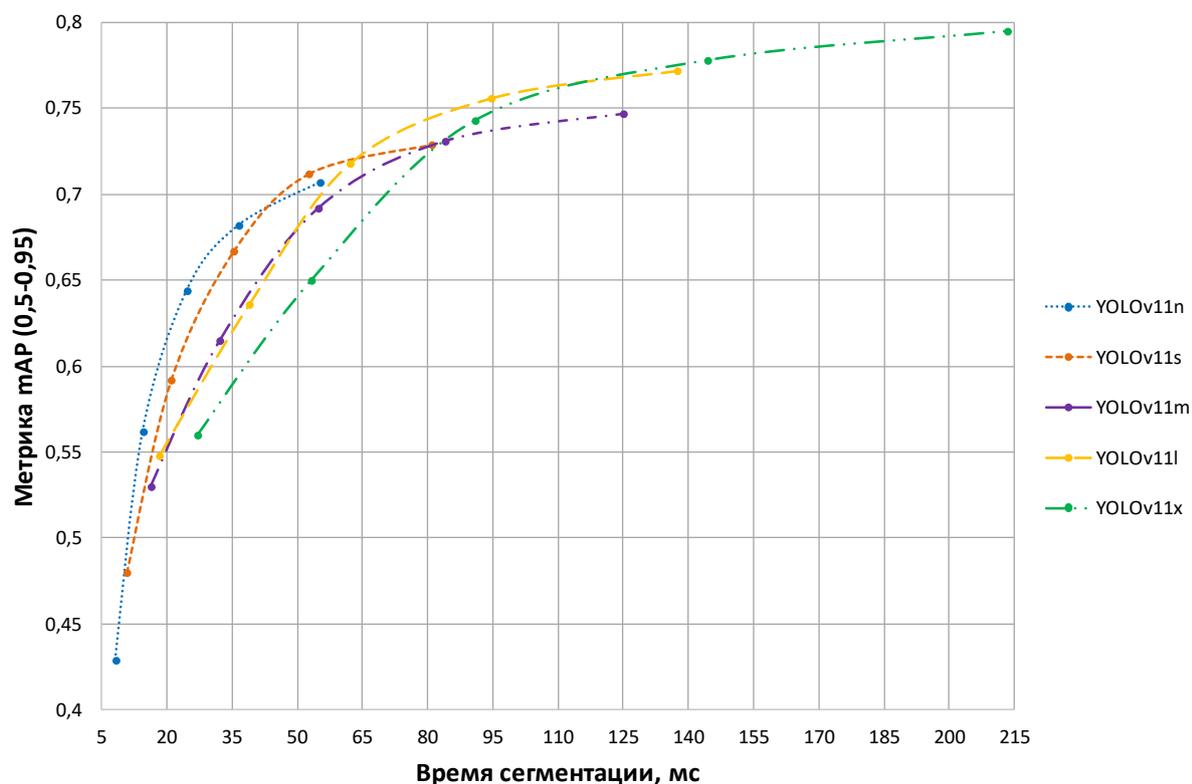


Рис. 4. График соотношения времени и точности сегментации

и усложнением версии модели. Наиболее быстрой и менее требовательной к мощности вычислительных ресурсов является версия YOLOv11n с разрешением входного слоя 640×640 пикселей.

Для комплексной сравнительной оценки эффективности различных конфигураций модели YOLOv11 с точки зрения компромисса между точностью сегментации и скоростью обработки на рис. 4 представлен график соотношения времени сегментации объектов и точности сегментации. Для каждой версии модели определена зависимость на основе экспериментальных данных, полученных при обработке изображений с разрешением входного слоя от 640×640 до 1920×1920 пикселей, что обеспечивает объективность результатов сравнительного анализа в широком диапазоне характеристик доступных вычислительных ресурсов. Каждая кривая на графике построена по пяти точкам, которые соответствуют значениям разрешения входного изображения, указанным в табл. 3. Движение вдоль кривой в направлении роста точности и времени сегментации отражает увеличение разрешения входного изображения.

Полученное поле параметров позволяет выбрать конфигурацию модели YOLOv11 с учетом требований АСУ ЖД ПС по качеству сегментации объектов и доступной мощности вычислительной платформы. Так, для систем реального времени с жесткими ограничениями по задержке может быть выбрана легкая версия (nano, small) с меньшим разрешением, в то время как для задач анализа архивных данных, где приоритетом является максимальная точность, оправдано применение более тяжелой версии (large, extra large) с высоким разрешением.

Для визуальной демонстрации качества сегментации объектов на рис. 5 приведены примеры сегментированных объектов на изображениях из тестовой выборки, которая не участвовала в процессе обучения, с использованием модели YOLOv11m при разрешении входного слоя 1920×1920 пикселей. Приведенные примеры подтверждают эффективность применения модели в различных условиях съемки, включая широкий диапазон освещенности и вариативности погодных явлений.

Анализ результатов обработки изображений подтверждает, что модель YOLOv11 успешно



Рис. 5. Сегментация объектов с использованием модели YOLOv11m и изображений с разрешением 1920×1920 пикселей

сегментирует и обнаруживает как крупные объекты (железнодорожная колея, платформы, тупиковые призмы), так и мелкие объекты железнодорожной инфраструктуры (пикетные столбики и карликовые светофоры). Ключевым преимуществом выбранной модели сверточной нейронной сети является ее высокая надежность при работе в условиях повышенной неопределенности. Вычислительные эксперименты подтвердили устойчивую сегментацию объектов при различных погодных явлениях (дождь, снег, туман), а также при значительных изменениях уровня освещенности в течение суток (от яркого солнечного освещения в полдень до сумрака в ночное время суток). Продemonстрированная устойчивость к изменяющимся внешним факторам является существенным достоинством модели YOLOv11 при использовании в реальных условиях эксплуатации грузового железнодорожного транспорта.

Заключение

Новым практическим результатом выполненной работы является созданный и апробированный набор визуальных данных для обучения сверточных нейронных сетей с учетом

особенностей инфраструктуры российских железных дорог и уникального разнообразия климатических условий.

На созданном наборе данных обучено 25 различных конфигураций модели YOLOv11, различающихся по версии архитектуры и по разрешению входного слоя. По результатам оценки качества обученных конфигураций YOLOv11 составлена таблица оценочных параметров, которая представляет зависимость точности сегментации и времени обработки от структурных параметров модели. Упрощение архитектуры модели и снижение разрешения входного изображения приводит к пропорциональному сокращению времени обработки, но одновременно снижает точность сегментации. Так, самая компактная версия YOLOv11n при разрешении 640×640 пикселей продемонстрировала время обработки всего 8,2 мс на видеокарте NVIDIA RTX A5000, но точность mAP (0,5–0,95) при этом составила лишь 0,429.

Полученные результаты могут быть использованы при разработке программно-аппаратного обеспечения бортовых систем технического зрения для АСУ ЖД ПС. Построенные графики соотношения времени и точности

сегментации объектов с разрешением изображений видеопотока позволяют выбрать масштабную конфигурацию модели в зависимости от конкретных требований к точности сегментации объектов и скорости обработки данных в режиме реального времени. ▲

Список источников

1. Охотников А. Л. Проекты развития систем автоматического управления движением поездов / А. Л. Охотников, И. А. Волкова // Интеллектуальные системы и технологии на транспорте. — 2023. — № 1(25). — С. 25–31.
2. Иванов В. Ф., Градусов А. Н. Алгоритм комплексирования сенсорных данных для задач автоматического управления подвижным составом / В. Ф. Иванов, А. Л. Охотников, А. Н. Градусов // Автоматика на транспорте. — 2024. — Т. 10. — № 4. — С. 360–371. DOI: 10.20295/2412-9186-2024-10-04-360-371.
3. Казанская Л. Ф. Беспилотные поезда при грузовых и пассажирских железнодорожных перевозках / Л. Ф. Казанская, П. В. Алпаева, А. А. Щипицына // Транспортное дело России. — 2023. — № 6. — С. 128–131.
4. Розенберг Е. Н. Комплексные решения по повышению пропускной способности железных дорог / Е. Н. Розенберг, И. Н. Розенберг, А. В. Озеров // Труды АО «НИИАС». — 2021. — Т. 1. — Вып. 11. — С. 32–47.
5. Алексеев А. С. Модель оценки показателей безопасности функционирования автоматизированных систем диспетчерского управления движением поездов / А. С. Алексеев, А. В. Горелик, В. Ю. Горелик, А. В. Истомин и др. // Автоматизация в промышленности. — 2025. — № 1. — С. 10–13.
6. Федоров В. А. Сверточные нейронные сети как инструмент обнаружения объектов железнодорожной инфраструктуры / В. А. Федоров, О. М. Огородникова // Автоматизация в промышленности. — 2024. — № 12. — С. 20–23. — DOI: 10.25728/avtprom.2024.12.04.
7. Fedorov V. A. AI-based train localization using railway infrastructure object detection / V. A. Fedorov // Transportation Development Research. — 2024. — Vol. 2. — Iss. 2. — Pp. 1–11. — DOI: 10.55121/tdr.v2i2.379.
8. Faghih-Roohi S. Deep convolutional neural networks for detection of rail surface defects / S. Faghih-Roohi, S. Hajizadeh, A. Nunez // 2016 International joint conference on neural networks (IJCNN). IEEE, July 2016. — Pp. 2584–2589. — DOI: 10.1109/IJCNN.2016.7727522.
9. Zhao Y. A review on rail defect detection systems based on wireless sensors / Y. Zhao, Z. Liu, D. Yi // Sensors. — 2022. — Vol. 22. — Iss. 17. — P. 6409. — DOI: 10.3390/s22176409.
10. Wei X. Railway track fastener defect detection based on image processing and deep learning techniques: A comparative study / X. Wei, Z. Yang, Y. Liu // Engineering Applications of Artificial Intelligence. — 2019. — Vol. 80. — Pp. 66–81. — DOI: 10.1016/j.engappai.2019.01.008.
11. Wu Y. Automatic detection of arbitrarily oriented fastener defect in high-speed railway / Y. Wu, Y. Qin, Y. Qian // Automation in Construction. — 2021. — Vol. 131. — DOI: 10.1016/j.autcon.2021.103913.
12. Сычугов А. Н. Применение нейронных сетей для распознавания объектов на железнодорожном транспорте / А. Н. Сычугов, В. Н. Михайчиков, М. В. Чернышов // Известия Петербургского университета путей сообщения. — 2023. — Т. 20. — Вып. 2. — С. 478–491. — DOI: 10.20295/1815-588X2023-2-478-491.
13. Fedorov V. A. Railway infrastructure instance segmentation based on convolutional neural networks / V. A. Fedorov // 2023 International Russian Automation Conference (RusAutoCon), IEEE, September 2023. — Pp. 443–447. — DOI: 10.1109/RusAutoCon58002.2023.10272908.
14. Ogorodnikova O. M., Method of ripe tomato detecting for a harvesting robot / O. M. Ogorodnikova, W. Ali // AIP Conference Proceedings, May 2019. — DOI: 10.1063/1.5134297.
15. Everingham M. The Pascal visual object classes (VOC) challenge / M. Everingham, L. Van Gool, C. K. Williams // International journal of computer vision. — 2010. — Vol. 88. — Pp. 303–338. — DOI: 10.1007/s11263-009-0275-4.

TRANSPORT AUTOMATION RESEARCH, 2025, Vol. 11, No. 3, pp. 239–249
DOI: 10.20295/2412-9186-2025-11-03-239-249

Object Segmentation Using Machine Vision in Automated Control Systems of Railway Rolling Stock

Information about authors

Fedorov V. A., Postgraduate Student. E-mail: Fedorov.Vladimir@UrFU.Ru
Ogorodnikova O. M., D. Sci. in Engineering, Associate Professor, Professor of the Department. E-mail: Olga.Ogorodnikova@UrFU.Ru

Ural Federal University, Department of Electronic Engineering, Yekaterinburg

Abstract: The paper presents the results of a study into the application of YOLOv11 convolutional neural network model configurations to object segmentation tasks. It is hypothesized that YOLOv11 will be utilized in the domain of machine vision within the context of an automated railway rolling stock control system. The primary focus of this study was to assess the performance and quality of image processing implemented by various model configurations. A comparative analysis of 25 YOLOv11 configurations was

conducted, varying in architecture and input image resolution within the range of 640×640 and 1920×1920 pixels. The creation of a specialized visual dataset was instrumental in facilitating the training of the models. This dataset comprised 20,000 annotated images of railway infrastructure, systematically distributed across 40 object classes. The performance and segmentation accuracy of all trained models was evaluated using the mAP metric (0.5–0.95). The results obtained are informative when choosing the configuration of the YOLOv11 model with the most suitable parameters for use in on-board machine vision systems of automated control systems, depending on the requirements for the accuracy of object segmentation and the power of available computing resources.

Keywords: machine vision; automated control system; railway rolling stock; object detection; object segmentation; convolutional neural network; YOLOv11.

References

1. Okhotnikov A. L., Volkova I. A. Proekty razvitiya sistem avtomaticheskogo upravleniya dvizheniem poezdov [Projects of development of automatic train control systems]. *Intellektual'nye sistemy i tekhnologii na transporte* [Intellectual systems and technologies in transport]. 2023, Iss. 1(25), pp. 25–31. (In Russian)

2. Ivanov V. F., Okhotnikov A. L., Gradusov A. N. Algoritm kompleksirovaniya sensorykh dannyykh dlya zadach avtomaticheskogo upravleniya podvizhnym sostavom [Algorithm for complexing sensor data for automatic rolling stock control tasks]. *Avtomatika na transporte* [Transport Automation Research]. 2024, vol. 10, Iss. 4, pp. 360–371. DOI: 10.20295/2412-9186-2024-10-04-360-371 (In Russian)
3. Kazanskaya L. F., Alpaeva P. V., Shchipitsyna A. A. Bepilotnye poezda pri gruzovykh i passazhirskikh zheleznodorozhnykh perevozkakh [Driverless trains in cargo and passenger railway transportation]. *Transportnoe delo Rossii* [Transport business of Russia]. 2023, Iss. 6, pp. 128–131. (In Russian)
4. Rozenberg E. N., Rozenberg I. N., Ozerov A. V. Kompleksnye resheniya po povysheniyu propusknoy sposobnosti zheleznykh dorog [Complex solutions to increase railway capacity]. *Trudy AO "NIIAS"* [Proceedings of JSC "NIIAS"]. 2021, vol. 1, Iss. 11, pp. 32–47. (In Russian)
5. Alekseev A. S., Gorelik A. V., Gorelik V. Yu., Istomin A. V. et al. Model' otsenki pokazateley bezopasnosti funkcionirovaniya avtomatizirovannykh sistem dispetcherskogo upravleniya dvizheniem poezdov [Model for assessing the safety indicators of automated train dispatching control systems]. *Avtomatizatsiya v promyshlennosti* [Automation in industry]. 2025, Iss. 1, pp. 10–13. (In Russian)
6. Fedorov V. A., Ogorodnikova O. M. Svertochnye neyronnye seti kak instrument obnaruzheniya ob'ektov zheleznodorozhnoy infrastruktury [Convolutional neural networks as a tool for detecting railway infrastructure objects]. *Avtomatizatsiya v promyshlennosti* [Automation in industry]. 2024, Iss. 12, pp. 20–23. DOI: 10.25728/avtprom.2024.12.04 (In Russian)
7. Fedorov V. A. AI-based train localization using railway infrastructure object detection. *Transportation Development Research*. 2024, vol. 2, Iss. 2, pp. 1–11. DOI: 10.55121/tdr.v2i2.379.
8. Faghih-Roohi S., Hajizadeh S., Nunez A. Deep convolutional neural networks for detection of rail surface defects. 2016 International joint conference on neural networks (IJCNN). IEEE, July 2016, pp. 2584–2589. DOI: 10.1109/IJCNN.2016.7727522.
9. Zhao Y., Liu Z., Yi D. A review on rail defect detection systems based on wireless sensors. *Sensors*. 2022, vol. 22, Iss. 17, p. 6409. DOI: 10.3390/s22176409.
10. Wei X., Yang Z., Liu Y. Railway track fastener defect detection based on image processing and deep learning techniques: A comparative study. *Engineering Applications of Artificial Intelligence*. 2019, vol. 80, pp. 66–81. DOI: 10.1016/j.engappai.2019.01.008.
11. Wu Y., Qin Y., Qian Y. Automatic detection of arbitrarily oriented fastener defect in high-speed railway. *Automation in Construction*. 2021, vol. 131. DOI: 10.1016/j.autcon.2021.103913.
12. Sychugov A. N., Mikheychikov V. N., Chernyshov M. V. Primenenie neyronnykh setey dlya raspoznavaniya ob'ektov na zheleznodorozhnom transporte [Application of neural networks for object recognition in railway transport]. *Izvestiya Peterburgskogo universiteta putey soobshcheniya* [Proceedings of Petersburg Transport University]. 2023, vol. 20, Iss. 2, pp. 478–491. DOI: 10.20295/1815-588X2023-2-478-491 (In Russian)
13. Fedorov V. A. Railway infrastructure instance segmentation based on convolutional neural networks. 2023 International Russian Automation Conference (RusAutoCon), IEEE, September 2023, pp. 443–447. DOI: 10.1109/RusAutoCon58002.2023.10272908
14. Ogorodnikova O. M., Ali W. Method of ripe tomato detecting for a harvesting robot. *AIP Conference Proceedings*, May 2019. DOI: 10.1063/1.5134297.
15. Everingham M., Van Gool L., Williams C. K. The Pascal visual object classes (VOC) challenge. *International journal of computer vision*. 2010, vol. 88, pp. 303–338. DOI: 10.1007/s11263-009-0275-4.