

# ВЫБОР АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЗАДАЧАХ ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ И ЛОКАЛИЗАЦИИ ОБЪЕКТОВ НА ЖЕЛЕЗНОЙ ДОРОГЕ

**БАРАНОВ Леонид Аврамович**, доктор техн. наук, профессор, профессор кафедры «Управление и защита информации»; e-mail: baranov.miit@gmail.com

**КУЛАГИН Максим Алексеевич**, кандидат техн. наук, доцент кафедры «Управление и защита информации»; e-mail: maksimkulagin06@yandex.ru

**ЯНЧЕНКО Герман Олегович**, аспирант кафедры «Управление и защита информации»; e-mail: ya@yanchenko.ru

Российский университет транспорта (МИИТ), Москва

Статья посвящена проблеме выбора и обоснования архитектур нейронных сетей для систем технического зрения на железнодорожном транспорте. В условиях цифровой трансформации отрасли важным становится создание надежных интеллектуальных систем поддержки принятия решений, способных работать в режиме реального времени. В работе проведен сравнительный анализ эффективности современных сверточных нейронных сетей и перспективных трансформерных архитектур в задаче семантической сегментации ключевых объектов железнодорожной инфраструктуры. Исследование базируется на специализированном наборе данных, включающем 8203 изображения, полученных с бортовых камер локомотивов в различных погодных и осветительных условиях. Протестированы пять моделей искусственного интеллекта: U-Net, U-Net++, DeepLabV3+, MAnet и SegFormer. Оценка эффективности проводилась по критериям, включающим метрику Mean Intersection over Union, время обработки одного кадра, а также анализ вероятностей ошибок первого (ложная тревога) и второго (пропуск объекта) рода. Экспериментально установлено, что архитектура U-Net++ с энкодером DenseNet-121 обеспечивает наиболее сбалансированные показатели, минимизируя критические пропуски сигналов при высокой точности локализации. Полученные результаты и предложенные критерии оценки позволяют обоснованно подходить к выбору алгоритмов машинного зрения при проектировании автономных систем управления и интеллектуальной системы поддержки принятия решений машинистов путевых машин.

**Ключевые слова:** семантическая сегментация, сверточные нейронные сети, трансформеры, техническое зрение, железнодорожная инфраструктура, искусственный интеллект

DOI: 10.20295/2412-9186-2026-12-02-128-137

## ▼ Введение

Современный этап развития транспортного комплекса Российской Федерации характеризуется масштабной цифровой трансформацией. Согласно стратегическим направлениям в области цифровой трансформации науки и высшего образования<sup>1</sup>, внедрение технологий искусственного интеллекта (ИИ) является приоритетным вектором развития высокотехнологичных отраслей. Внедрение цифровых технологий и развитие интеллектуальных транспортных систем (ИТС) являются ключевыми направлениями модернизации железно-

дорожной отрасли. В современной литературе уделяется значительное внимание вопросам автоматизации технологических процессов [1] и роботизации инфраструктуры [2], а также перспективам перехода к автономному вождению [3] и внедрению комплексов автоматизированного управления движением поездов [4].

Одним из основным элементов таких интеллектуальных систем выступают технологии машинного зрения, позволяющие эффективно решать задачи сбора и анализа данных, оценки состояния пути и распознавания объектов [5, 6]. При этом широкое внедрение подобных решений требует особого внимания к вопросам обеспечения функциональной и информационной безопасности: исследователи активно прорабатывают методы защиты

<sup>1</sup> Об утверждении стратегического направления в области цифровой трансформации науки и высшего образования: распоряжение Правительства Российской Федерации от 21.12.2021 № 3759-р.

автоматизированных систем управления и объектов критической информационной инфраструктуры (КИИ) от угроз [7, 8], в том числе с учетом специфики алгоритмов машинного обучения [9]. Также известны работы по созданию и внедрению систем автоматического управления движением поездов внеуличного городского транспорта, магистральных железных дорог, в том числе высокоскоростных пассажирских поездов, развивается новое направление автоматического управления путевых машин [10, 11].

Аппаратные решения и алгоритмы функционирования систем управления зависят от уровня автоматизации [12], требований к безопасности движения, диапазона скоростей и условий эксплуатации. Одним из требований, предъявляемых к автоматическим системам управления подвижным составом, является определение свободности пути следования подвижного состава, что делает необходимым использовать системы технического зрения (СТЗ). Соответствующие исследования ведутся в ряде организаций: РУТ (МИИТ), АО «ВНИИЖТ», АО «НИИАС». Аналогичными разработками занимаются зарубежные компании (Siemens Mobility, Alstom, Thales, Stadler) [13].

Концепция построения подсистемы технического зрения для интеллектуальной системы поддержки принятия решений (ИСППР) для автоматического управления путевых машин рассмотрена в [10]. В рамках данной статьи описывается задача семантической сегментации, решение которой осуществляется с помощью моделей ИИ. В большинстве существующих исследований для сегментации железнодорожной инфраструктуры применяются сверточные нейронные сети (CNN) [14, 15]. Развитие ИИ привело к появлению новых архитектур нейронных сетей (НС), использующих трансформеры (transformers). В отличие от сверточных НС, работающих с локальными признаками, трансформерные НС используют механизмы глобального внимания (global attention), что позволяет эффективнее учитывать контекст всей сцены при попиксельной классификации. Исследование эффективности такого подхода в условиях управления движением подвижных единиц железнодорожного транспорта является актуальным.

Для оценки качества функционирования подсистем технического зрения используются два следующих статистических критерия:

- ошибка первого рода ( $\alpha$ ) — вероятность того, что препятствие находится на пути при его фактическом отсутствии (ложная тревога);
- ошибка второго рода ( $\beta$ ) — вероятность сигнала об отсутствии препятствия при фактическом его наличии (пропуск сигнала).

Целью данной работы является научно обоснованный выбор и верификация архитектуры НС для интеграции в подсистему технического зрения ИСППР на железнодорожном транспорте.

Для достижения поставленной цели сформулированы и решаются следующие задачи:

1. Подобрать из открытых источников специализированный набор данных, содержащий изображения с бортовых камер локомотивов в различных погодных условиях и при разной освещенности.
2. Реализовать и обучить современные сверточные (U-Net, U-Net++, DeepLabV3+, MAnet) и трансформерные (SegFormer) архитектуры НС для задачи семантической сегментации объектов железнодорожного пути.
3. Провести сравнительный анализ эффективности исследуемых моделей ИИ по комплексному критерию, включающему Mean Intersection over Union (mIoU), время обработки одного кадра, вероятности ошибок первого и второго рода.
4. Выявить наиболее сбалансированную архитектуру НС, обеспечивающую минимальный уровень критических ошибок при сохранении требуемого быстродействия в режиме реального времени.

### **Методология решения задачи, используемый инструментарий**

Решение задачи семантической сегментации в НС базируется на архитектурном шаблоне «энкодер-декодер» (encoder-decoder), который позволяет преобразовывать входное изображение в попиксельную карту классов.

- Энкодер (encoder) — это часть нейронной сети, отвечающая за блок формирования признаков (backbone). Его задача заключается в постепенном сжатии входного

изображения и извлечении из него высокоуровневой семантической информации. В процессе обработки в энкодере происходит уменьшение пространственного разрешения (размера картинка) и увеличение количества каналов, что позволяет сети «понимать», какие именно объекты находятся в кадре (например, рельсы или поезд), игнорируя фоновые шумы.

- Декодер (decoder) — это восстанавливающая часть сети. Его задача — перенести извлеченные энкодером признаки обратно в исходное пространственное разрешение изображения. Декодер сопоставляет накопленную семантическую информацию с конкретными координатами пикселей, что позволяет точно локализовать границы объектов инфраструктуры.

В данном исследовании в качестве энкодеров рассматриваются следующие семейства архитектур: ResNet [16], EfficientNet [17] и DenseNet [18]. Применение НС, предварительно обученных на больших наборах данных, реализует механизм обучения (transfer learning), что позволяет ускорить процесс настройки моделей ИИ.

Сравнительный анализ эффективности попиксельной классификации проводился для сверточных нейронных сетей U-Net [19], U-Net++ [20], DeepLabV3+ [21] и MAnet [22]. Наряду с классическими сверточными решениями в данной работе рассматривается современная трансформерная архитектура SegFormer [23]. Выбранный набор моделей ИИ позволяет оценить эффективность различных архитектурных решений в задачах семантической сегментации для СТЗ на железнодорожном транспорте.

Для проведения имитационного моделирования функционирования моделей ИИ использовался программный набор на базе языка программирования Python и библиотеки глубокого обучения PyTorch. Проектирование и обучение моделей семантической сегментации осуществлялось с применением открытой библиотеки segmentation\_models.pytorch, которая предоставляет унифицированный интерфейс для интеграции различных НС и архитектурных компонентов. Все расчеты производились с использованием программно-аппаратной архитектуры CUDA на графическом ускорителе NVIDIA

RTX A5000 с вычислительной мощностью по количеству операций в секунду 27,8 TFLOPS и объемом видеопамати (VRAM) 24 Гб.

### Описание данных и подготовка эксперимента

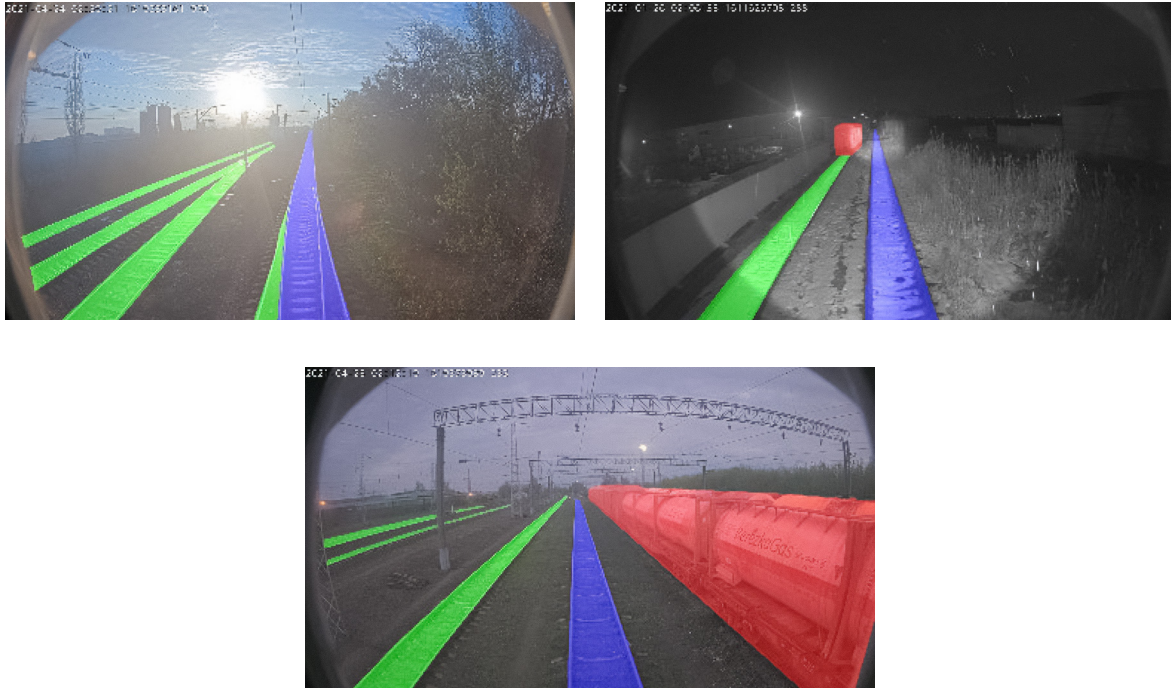
Для обучения и тестирования моделей ИИ использовался специализированный набор данных RailDataSets от Санкт-Петербургского филиала АО «НИИАС» [24], сформированный на основе видеопотоков с бортовых камер локомотивов, работающих в реальных условиях эксплуатации. Общий объем выборки составил 8203 изображения.

Данные для проведения имитационных экспериментов содержат попиксельную разметку для трех классов объектов, значимых для функционирования СТЗ и ИСППР:

1. Путь движения подвижного состава (MainRailPolygon) — область рельсового полотна, по которому движется локомотив.
2. Соседние пути движения (AlternativeRailPolygon) — смежные железнодорожные пути.
3. Поезда (Train) — вагоны и локомотивы, находящиеся в зоне видимости.

В рамках предобработки данных все исходные изображения приводились к единому разрешению  $512 \times 512$  пикселей. Выбор данного геометрического размера входного слоя продиктован необходимостью последующей экспериментальной проверки качества функционирования ИСППР с применением реального бортового вычислительного устройства. Его специфика заключается в ограничениях по объему VRAM, который распределяется не только на модель семантической сегментации, но и на другие задачи, решаемые системой. Разрешение  $512 \times 512$  пикселей является максимально возможным для используемого бортового вычислительного устройства, что гарантирует обработку видеопотока в режиме реального времени без превышения лимитов памяти и возникновения ошибок нехватки ресурсов.

Для увеличения числа обучающей выборки проведена аугментация исходных данных: случайное изменение яркости и контрастности, а также внесение геометрических искажений. Выборка была разделена в пропорции: 80 % — обучающая, 10 % — валидационная и 10 % — тестовая.



**Рис. 1.** Примеры изображений из набора данных с попиксельной разметкой (синий — MainRailPolygon, зеленый — AlternativeRailPolygon, красный — Train)

### Имитационные эксперименты и их результаты

Оценка критериев качества моделей семантической сегментации проводилась с использованием метрики Intersection over Union (IoU), также известной как индекс Жаккара. Данная метрика определяет степень перекрытия предсказанной маски ( $A_i$ ) и эталонной разметки ( $B_i$ ):

$$IoU_i = \frac{|A_i \cap B_i|}{|A_i \cup B_i|}, \quad (1)$$

где  $i$  — номер класса.

Для оценки качества функционирования моделей ИИ по всем классам объектов, опи-

санных выше, использовалось усреднение значения метрики IoU (2):

$$mIoU = \frac{1}{N} \sum_{i=1}^N IoU_i, \quad (2)$$

где  $N$  — число классов.

Для оценки влияния архитектуры энкодера на характеристики моделей был проведен анализ по единой методике для всех исследуемых архитектур. В качестве примера в табл. 1 представлены результаты для архитектуры U-Net++, демонстрирующие зависимость качества сегментации, времени обработки одного кадра и объема модели (количество обучаемых параметров) от выбора энкодера.

**ТАБЛИЦА 1.** Влияние выбора энкодера на характеристики архитектуры U-Net++

Энкодер	Параметры, млн	Время обработки одного кадра, мс	IoU MainRail	IoU AlternativeRail	IoU Train	mIoU
ResNet-34	26	9,39	0,873	0,685	0,763	0,774
ResNet-101	68	20,13	0,861	0,699	0,773	0,778
EfficientNet-b2	10	20,05	0,868	0,674	0,763	0,768
EfficientNet-b4	21	32,67	0,887	0,699	0,790	0,792
DenseNet-121	30	31,51	0,898	0,721	0,796	0,805
DenseNet-201	48	38,83	0,901	0,716	0,797	0,805

Аналогичный анализ был выполнен для всех архитектур. Для DeepLabV3+ тестирование с энкодерами семейства DenseNet технически не осуществлялось ввиду ограничений совместимости в используемой реализации библиотеки `segmentation_models.pytorch`. Для трансформерной архитектуры SegFormer применялись специализированные энкодеры на базе MiT (Mix Transformer) [23].

По результатам, представленным в табл. 1, можно сделать следующие выводы:

1. Рост числа параметров энкодера не гарантирует пропорционального улучшения качества: EfficientNet-b2 (10 млн параметров) показывает  $mIoU = 0,768$ , уступая ResNet-34 (26 млн,  $mIoU = 0,774$ ).

2. Архитектуры на базе DenseNet демонстрируют наилучшее качество сегментации.

3. Увеличение глубины энкодера (ResNet-34 → ResNet-101) дает незначительный прирост  $mIoU (+0,004)$  при росте времени обработки кадра более чем в 2 раза.

4. Лучшей комбинацией для архитектуры U-Net++ является связка с DenseNet-121.

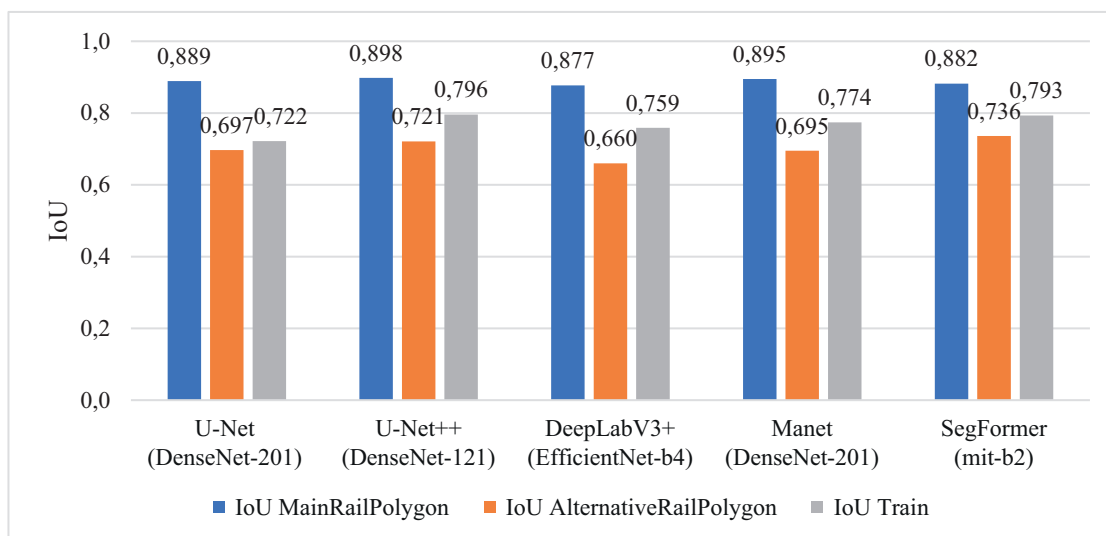
По описанной методике был проведен анализ влияния выбора энкодера для всех пяти исследуемых архитектур. Сводные результаты, отражающие лучшие конфигурации каждой архитектуры по критериям, представлены в табл. 2 и 3.

Для детальной оценки качества сегментации проведен анализ значений критерия IoU по отдельным классам объектов. На рис. 2 представлены результаты для конфигураций моделей, отобранных по критерию  $mIoU$  (табл. 2).

На рис. 2 представлены значения IoU по классам объектов для каждого типа архитектуры моделей с наилучшими значениями критерия  $mIoU$ . Согласно графику, наименьшее качество сегментации зафиксировано для класса `AlternativeRailPolygon` (0,66–0,74), что свидетельствует о его наибольшей сложности для распознавания всеми

**ТАБЛИЦА 2. Лучшие конфигурации каждой архитектуры по критерию качества ( $mIoU$ )**

Архитектура	Энкодер	Параметры, млн	$mIoU$
U-Net	DenseNet-201	28	0,769
U-Net++	DenseNet-121	30	0,805
DeepLabV3+	EfficientNet-b4	18	0,765
MANet	DenseNet-201	133	0,791
SegFormer	mit-b2	27	0,804



**Рис. 2.** Значения критерия качества сегментации по классам объектов для исследуемых архитектур

**ТАБЛИЦА 3. Лучшие конфигурации каждой архитектуры по критерию времени обработки одного кадра**

Архитектура	Энкодер	Параметры, млн	Время обработки одного кадра, мс	mIoU
U-Net	ResNet-34	24	8,10	0,764
U-Net++	ResNet-34	26	9,39	0,774
DeepLabV3+	ResNet-34	22	7,02	0,753
MAnet	ResNet-34	32	11,00	0,764
SegFormer	mit-b2	27	21,48	0,804

архитектурами. Остальные классы, в частности MainRailPolygon и Train, определяются моделью увереннее.

Помимо качества сегментации, важным требованием к моделям ИИ в составе бортовых СТЗ является скорость обработки одного кадра, определяющая возможность работы с видеопотоком в режиме реального времени. В табл. 3 представлены лучшие конфигурации каждой архитектуры, отобранные по минимальному времени обработки одного кадра. Данные результаты позволяют оценить компромисс между точностью и временем обработки одного кадра для использования моделей на бортовых вычислительных устройствах.

Анализ данных, представленных в табл. 3, показывает, что использование энкодера ResNet-34 позволяет достичь минимального времени обработки одного кадра (7–11 мс) для всех сверточных архитектур. Наилучшей по времени обработки одного кадра является модель DeepLabV3+ с энкодером ResNet-34 (7,02 мс), что делает данную конфигурацию предпочтительной для систем с жесткими требованиями к времени работы. Архитектура SegFormer при времени обработки одного кадра 21,48 сохраняет качество сегментации на уровне лидера по точности сегментации (mIoU = 0,804).

По данным из табл. 2 и 3 можно сделать следующие выводы:

1. Лидерами по качеству сегментации оказались модели архитектуры Unet++, SegFormer и MaNet. При использовании энкодеров типа DenseNet наблюдается наилучшее качество работы, однако время обработки одного кадра возрастает по сравнению с другими рассмотренными энкодерами.

2. Повышение числа обучаемых параметров у энкодеров приводит к увеличению качества работы моделей ИИ, но иногда это увеличение незначительно. Время обработки одного кадра при этом может возрастать в 1,5–2 раза.

3. По полученным данным сверточная модель архитектуры U-Net++ с энкодером DenseNet-121 показала себя лидером по критерию mIoU (0,805), в то же время трансформерная модель SegFormer с энкодером mit-b2 демонстрирует практически идентичное значение mIoU (0,804). Модель SegFormer (mit-b2) достигает таких значений при меньшем количестве обучаемых параметров (27 млн), чем U-Net++ (DenseNet-121) с 30 млн параметров. Модель U-Net++ (DenseNet-121) обрабатывает один кадр в среднем за 31,51 мс, у модели SegFormer время обработки составляет 21,48 мс.

Анализ результатов, представленных выше, позволяет перейти от оценки качества сегментации к оценке применимости моделей ИИ в составе бортовых СТЗ железнодорожного подвижного состава. Используя данные о качестве сегментации (mIoU) моделей и времени обработки одного кадра, для дальнейшего исследования в составе ИСППР были отобраны три архитектуры:

1. U-Net++ (backbone: DenseNet-121) выбрана как архитектура, показавшая наивысшие критерии качества.

2. SegFormer (backbone: mit-b2) включена в анализ как современная альтернатива классическим сверточным сетям на базе трансформеров. Модель продемонстрировала сопоставимый с лидером уровень качества.

3. DeepLabV3+ (backbone: Resnet-34) выбрана как наиболее быстродействующая модель из рассмотренных.

### Методология оценки эксплуатационной надежности в составе ИСППР

В рамках разработки ИСППР машиниста специальной путевой машины [10] критерии качества семантической сегментации (IoU, mIoU) недостаточно для полной оценки безопасности системы. Учитывая известные трудности формирования матрицы штрафов, в качестве характеристик, существенно определяющих качество работы системы, будем считать вероятности  $\alpha$ - и  $\beta$ -ошибок.

Для расчета вероятностей ошибок была собрана тестовая выборка из 1000 изображений, содержащих препятствия на пути. Используемый алгоритм обнаружения препятствий является предварительным и служит только для сравнения архитектур сегментации в одинаковых условиях. Все модели тестировались при одинаковых настройках обработки. Показатели вероятностей ошибок, приведенные в табл. 4, достаточно высокие, так как разработка всей системы находится на начальном этапе и не проводилась полная настройка всех ее компонентов. Тем не менее разница в результатах между моделями позволяет объективно сравнить их эксплуатационную надежность. В табл. 4 приведены расчетные значения вероятностей возникновения  $\alpha$ - и  $\beta$ -ошибок для отобранных архитектур. Эти данные позволяют оценить не только качество сегментации, но и потенциальные риски при эксплуатации моделей ИИ в составе систем безопасности.

**ТАБЛИЦА 4. Ошибки первого и второго рода для отобранных моделей**

Архитектура	$\alpha$ -ошибка	$\beta$ -ошибка
U-Net++	0,36	0,37
DeepLabV3+	0,29	0,39
SegFormer	0,20	0,43

Анализ результатов, представленных в табл. 4, дает возможность определить перспективы внедрения конкретных ИИ-моделей в состав ИСППР:

1. Архитектура U-Net++ показала себя как наиболее надежное решение с точки зрения минимизации  $\beta$ -ошибки. Поскольку в железнодорожной сфере пропуск объекта на пути

представляет наибольшую угрозу, данная модель является приоритетной для алгоритмов обнаружения препятствия.

2. Модель SegFormer продемонстрировала самую низкую вероятность  $\alpha$ -ошибки, что означает минимальное число ложных срабатываний. Данное свойство важно для систем мониторинга, целью которых является исключение необоснованных задержек в движении поездов. Тем не менее риск пропуска сигнала у данной архитектуры остается значительным.

### Заключение

Проведенное исследование посвящено сравнительному анализу архитектур сверточных и трансформерных нейронных сетей применительно к задаче локализации элементов железнодорожной инфраструктуры. Полученные результаты свидетельствуют о том, что при близких показателях качества различных моделей определяющую роль в выборе архитектуры играют конкретные условия эксплуатации и состав требований к системе.

По совокупности критериев, связанных с обеспечением безопасности движения, наиболее перспективной признана связка U-Net++ с энкодером DenseNet-121. Данная конфигурация отличается не только высоким качеством семантической сегментации, но и минимальным среди рассмотренных вариантов уровнем пропусков целевых объектов. Поскольку для бортовых систем предупреждения столкновений снижение вероятности ошибки второго рода является приоритетным, именно эта архитектура рекомендуется к внедрению в ИСППР.

Там, где критичным требованием выступает минимизация ложных срабатываний, предпочтительным решением является SegFormer. Экспериментально установлено, что трансформерная архитектура демонстрирует наименьший уровень  $\alpha$ -ошибки. Сокращение числа ложных срабатываний исключает неоправданные торможения подвижного состава и связанные с ними потери пропускной способности.

В сценариях с жесткими ограничениями по вычислительным ресурсам и времени отклика обоснованным выбором является DeepLabV3+. Время обработки одного кадра

составляет 7,02 мс, что существенно превосходит остальные архитектуры по быстродействию. Показатели сегментации при этом остаются на приемлемом уровне, что делает модель пригодной для применения в системах реального времени.

Таким образом, универсального решения для всех условий эксплуатации не существует, а значит, выбор архитектуры НС должен опираться на комплексный анализ точности локализации, вероятностей ошибок первого и второго рода, а также производительности. Предложенная методика многокритериальной оценки обеспечивает обоснованный подбор алгоритмов под конкретные требования бортового оборудования. ▲

*Работа выполнена в рамках государственного задания от 15.01.2026 № 103-00001-26-00 за счет бюджетного финансирования.*

#### Список источников

1. Зелова М.И., Комаров А.В. О применении комплекса автоматизированного управления движением поездов в условиях высокой интенсивности движения // Современные технологии. Системный анализ. Моделирование. 2022. № 4 (76). С. 92–100. DOI: 10.26731/1813-9108.2022.4(76).92-100. EDN WFQNPA
2. Роботизация на железнодорожном транспорте / А.И. Долгий [и др.] // Интеллектуальный транспорт. 2025. Т. 9, № 3 (35). С. 4–32. EDN JCBFIE
3. Шебе Х., Шубинский И.Б., Розенберг Е.Н. Различные подходы к автономному вождению для железных дорог // Надежность. 2025. Т. 25, № 1. С. 4–10. DOI: 10.21683/1729-2646-2025-25-1-4-10. EDN DTBFTK
4. Корнев Д.А. Моделирование сетей передачи данных в интеллектуальной системе управления движением поездов // Известия Транссиба. 2023. № 3 (55). С. 141–154. EDN QICJOC
5. Осипов Д.В. Машинное зрение на транспорте: перспективы и применение // Экономика железных дорог. 2023. № 8. С. 71–75. EDN IAHQWG
6. Тимофеев Е.Н., Ефимов Н.В., Комягин С.А. Подготовка данных для создания цифрового проекта на участок ремонта железнодорожного пути с использованием САУ-ЗД // Интерэкспо Гео-Сибирь. 2024. Т. 1. С. 116–121. DOI: 10.33764/2618-981X-2024-1-116-121. EDN FQOUYJ
7. Баранов Л.А., Иванова Н.Д., Михалевич И.Ф. Цифровой испытательный стенд анализа безопасности объектов критической информационной инфраструктуры интеллектуальных систем водного транспорта // Надежность. 2025. Т. 25, № 3. С. 50–59. DOI: 10.21683/1729-2646-2025-25-3-50-59. EDN QGPPKI
8. Комплексная безопасность АСУ ТП объектов КИИ железнодорожного транспорта / П.А. Попов [и др.] // Надежность. 2024. Т. 24, № 4. С. 48–57. DOI: 10.21683/1729-2646-2024-24-4-48-57. EDN YKHBOB
9. Баранов Л.А., Сафронов А.И., Сидоренко В.Г. Развитие интеллектуальных систем управления электрическим транспортом // Автоматика, связь, информатика. 2025. № 10. С. 30–32. DOI: 10.62994/AT.2025.10.10.007. EDN BOSBZK
10. Баранов Л.А., Кулагин М.А., Янченко Г.О. Техническое зрение в интеллектуальной системе поддержки принятия решений машиниста специальной путевой машины // Наука и техника транспорта. 2025. № 3. С. 65–69. EDN PZPESY
11. Кулагин М.А., Янченко Г.О. Подходы к определению препятствий на железнодорожном пути с помощью нейронных сетей // Интеллектуальные транспортные системы: материалы III Международной научно-практической конференции (Москва, 30 мая 2024 года). М.: Российский университет транспорта (МИИТ), 2024. С. 626–630. DOI: 10.30932/9785002446094-2024-626-630. EDN OUMPCR
12. Никulichиков П.М. История, состояние и перспективы развития систем автоматического управления поездами метрополитена // Автоматика на транспорте. 2016. Т. 2, № 3. С. 456–473. EDN WYXBNL
13. Озеров А.В., Бочков А.В. Текущее состояние и перспективы мировых исследований в области интеллектуального железнодорожного транспорта // Интеллектуальный транспорт. 2025. Т. 9, № 2 (34). С. 43–66. EDN QFUBXM
14. Медведева Е.В., Перевощикова А.А. Исследование нейросетевых алгоритмов распознавания объектов железнодорожной инфраструктуры на видеозображениях // Компьютерная оптика. 2025. Т. 49, № 3. С. 443–450. DOI: 10.18287/2412-6179-CO-1563. EDN JUAWIT
15. Федоров В.А., Огородникова О.М. Сегментация объектов техническим зрением в автоматизированных системах управления железнодорожным подвижным составом // Автоматика на транспорте. 2025. Т. 11, № 3. С. 239–249. DOI: 10.20295/2412-9186-2025-11-03-239-249. EDN IUNEIX
16. Deep Residual Learning for Image Recognition / K. He [et al.] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016 DOI: 10.1109/CVPR.2016.90

17. Tan M., Le Q.V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // International Conference on Machine Learning (ICML). 2019. DOI: 10.48550/arXiv.1905.11946
18. Densely Connected Convolutional Networks / G. Huang [et al.] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. DOI: 10.1109/CVPR.2017.243
19. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015. DOI: 10.1007/978-3-319-24574-4\_28
20. UNet++: A Nested U-Net Architecture for Medical Image Segmentation / Z. Zhou [et al.] // International Workshop on Deep Learning in Medical Image Analysis (DLMIA/MICCAI). 2018. DOI: 10.1007/978-3-030-00889-5\_1
21. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation / Chen L.-C. [et al.] // European Conference on Computer Vision (ECCV), 2018. DOI: 10.48550/arXiv.1802.02611
22. Multi-Attention-Network for Semantic Segmentation of Fine Resolution Remote Sensing Images / R. Li [et al.] // IEEE Transactions on Geoscience and Remote Sensing (TGRS). 2022. DOI: 10.1109/TGRS.2021.3093977
23. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers / E. Xie [et al.] // Conference on Neural Information Processing Systems (NeurIPS). 2021. DOI: 10.48550/arXiv.2105.15203
24. Popov P.A., Ozerov A.V., Marshova A.S. Prospects of autonomous railway transport development, BRICS Transport. 2024. № 3 (3). Pp. 1–14. DOI: 10.46684/2024.3.4. EDN: HPYXEZ

TRANSPORT AUTOMATION RESEARCH. 2026. Vol. 12, no. 2, pp. 128–137  
DOI: 10.20295/2412-9186-2026-12-02-128-137

### Selection of Artificial Intelligence Algorithms in Computer Vision Tasks for Detection and Localization of Objects on Railways

#### Information about authors

**Baranov L. A.**, Dr. Sci. in Engineering, Professor, Professor of the Department “Management and Information Security”; e-mail: baranov.miit@gmail.com

**Kulagin M. A.**, PhD in Engineering, Associate Professor of the Department “Management and Information Security”; e-mail: maksimkulagin06@yandex.ru

**Yanchenko G. O.**, Postgraduate Student of the Department “Management and Information Security”; e-mail: ya@gyanchenko.ru

Russian University of Transport (MIIT), Moscow

**Abstract:** this article addresses the problem of selecting and substantiating neural network architectures for computer vision systems in railway transport. In the context of digital transformation within the industry, the development of reliable intelligent decision support systems capable of real-time operation has become a critical priority. This study presents a comparative analysis of the effectiveness of modern convolutional neural networks and emerging transformer-based architectures for the semantic segmentation of key railway infrastructure objects. The research is based on a specialized dataset comprising 8,203 images captured by onboard locomotive cameras under diverse weather and lighting conditions. Five artificial intelligence models were evaluated: *U-Net*, *U-Net++*, *DeepLabV3+*, *MAnet*, and *SegFormer*. Performance assessment was conducted using criteria including the Mean Intersection over Union (mIoU) metric, inference speed, and analysis of Type I (false alarm) and Type II (missed detection) error probabilities. Experimental results demonstrate that the *U-Net++* architecture with a *DenseNet-121* encoder delivers the most balanced performance, minimizing critical missed detections while maintaining high localization accuracy. The findings and proposed evaluation criteria provide a structured framework for the substantiated selection of machine vision algorithms in the design of autonomous control systems and intelligent decision support systems for operators of track maintenance machinery.

**Keywords:** semantic segmentation, convolutional neural networks, transformers, computer vision, railway infrastructure, artificial intelligence

#### References

1. Zelova M. I., Komarov A. V. O primeneni kompleksa avtomatizirovannogo upravleniya dvizheniem poezdov v usloviyakh vysokoi intensivnosti dvizheniya [On the Application of a Complex of Automated Train Traffic Control in Conditions of High Traffic Intensity], *Sovremennye tekhnologii. Sistemnyi analiz. Modelirovanie* [Modern technologies. System Analysis. Simulation], 2022, no. 4 (76), pp. 92–100. DOI: 10.26731/1813-9108.2022.4(76).92-100. EDN WFNQNA (In Russian)
2. Dolgii A. I., et al. Robotizatsiya na zheleznodorozhnom transporte [Robotization in Railway Transport], *Intellektual'nyi transport* [Intelligent Transport], 2025, vol. 9, no. 3 (35), pp. 4–32. EDN JCBFIE (In Russian)
3. Shebe Kh., Shubinskii I. B., Rozenberg E. N. Razlichnye podkhody k avtonomnomu vozhdniyu dlya zheleznykh dorog [Different Approaches to Autonomous Driving for Railways], *Nadezhnost' [Reliability]*, 2025, vol. 25, no. 1, pp. 4–10. DOI: 10.21683/1729-2646-2025-25-1-4-10. EDN DTBFTK (In Russian)
4. Kornev D. A. Modelirovanie setei peredachi dannykh v intellektual'noi sisteme upravleniya dvizheniem poezdov [Modeling of Data Transmission Networks in an Intelligent Train Traffic Management System], *Izvestiya Transsiba* [Journal of Transsib Railway Studies], 2023, no. 3 (55), pp. 141–154. EDN QICJOC (In Russian)
5. Osipov D. V. Mashinnoe zrenie na transporte: perspektivy i primeneniye [Machine Vision in Transport: Prospects and Applications], *Ekonomika zheleznykh dorog* [Railway Economy], 2023, no. 8, pp. 71–75. EDN IAHQWG (In Russian)
6. Timofeev E. N., Efimov N. V., Komyagin S. A. Podgotovka dannykh dlya sozdaniya tsifrovogo proekta na uchastok remonta zheleznodorozhnogo puti s ispol'zovaniem SAU-3D [Data Preparation for the Creation of a Digital Project for the Railway Track Repair Site Using a 3D Automated Control System], *Interexpo Geo-Sibir* [Interexpo GEO-Siberia], 2024, vol. 1, pp. 116–121. DOI: 10.33764/2618-981X-2024-1-116-121. EDN FQOYUJ (In Russian)
7. Baranov L. A., Ivanova N. D., Mikhalevich I. F. Tsifrovoy ispytatel'nyi stend analiza bezopasnosti ob'ektov kriticheskoi informatsionnoi infrastruktury intellektual'nykh sistem vodnogo transporta [Digital Test Bench for Analyzing the Safety of Critical Information Infrastructure Facilities of Intelligent Water Transport Systems], *Nadezhnost' [Reliability]*, 2025, vol. 25, no. 3, pp. 50–59. DOI: 10.21683/1729-2646-2025-25-3-50-59. EDN QGPPKI (In Russian)
8. Popov P. A., et al. Kompleksnaya bezopasnost' ASU TP ob'ektov KII zheleznodorozhnogo transporta [Comprehensive Security of an Automated Process Control System for Critical Information Infrastructure], *Nadezhnost' [Reliability]*, 2024, vol. 24, no. 4, pp. 48–57. DOI: 10.21683/1729-2646-2024-24-4-48-57. EDN YKHBOB (In Russian)

9. Baranov L. A., Safronov A. I., Sidorenko V. G. Razvitie intellektual'nykh sistem upravleniya elektricheskim transportom [Development of Intelligent Electric Transport Management Systems], *Avtomatika, svyaz', informatika [Automation, Communications, Computer Science]*, 2025, no. 10, pp. 30–32. DOI: 10.62994/AT.2025.10.10.007. EDN BOSBZK (In Russian)
10. Baranov L. A., Kulagin M. A., Yanchenko G. O. Tekhnicheskoe zrenie v intellektual'noi sisteme podderzhki prinyatiya reshenii mashinista spetsial'noi putevoi mashiny [Technical Vision in the Intelligent Decision Support System of the Driver of a Special Track Vehicle], *Nauka i tekhnika transporta [Science and Technology of Transport]*, 2025, no. 3, pp. 65–69. EDN PZPESY (In Russian)
11. Kulagin M. A., Yanchenko G. O. Podkhody k opredeleniyu prepyatstviy na zheleznodorozhnom puti s pomoshch'yu neironnykh setei [Approaches to Determining Obstacles on a Railway Track Using Neural Networks], *Intellektual'nye transportnye sistemy: Materialy III Mezhdunarodnoi nauchno-prakticheskoi konferentsii [Intelligent Transport Systems: Proceedings of the III International Scientific and Practical Conference]*, May 30, 2024. Moscow, Russian University of Transport, 2024, pp. 626–630. DOI: 10.30932/9785002446094-2024-626-630. EDN OUMPCR (In Russian)
12. Nikul'chikov P. M. Istoriya, sostoyanie i perspektivy razvitiya sistem avtomaticheskogo upravleniya poezdami metropolitena [History, Status and Prospects of Development of Automatic Train Control Systems], *Avtomatika na transporte [Transport Automation Research]*, 2016, vol. 2, no. 3, pp. 456–473. EDN WYXBNL (In Russian)
13. Ozerov A. V. Bochkov A. V. Tekushchee sostoyanie i perspektivy mirovykh issledovaniy v oblasti intellektual'nogo zheleznodorozhnogo transporta [The Current State and Prospects of Global Research in the Field of Intelligent Railway Transport], *Intellektual'nyi transport [Intelligent Transport]*, 2025, vol. 9, no. 2 (34), pp. 43–66. EDN QFUBXM (In Russian)
14. Medvedeva E. V., Perevoshchikova A. A. Issledovanie neirosetevykh algoritmov raspoznavaniya ob"ektov zheleznodorozhnoi infrastruktury na videoizobrazheniyakh [Investigation of Neural Network Algorithms for Recognizing Railway Infrastructure Objects in Video Images], *Komp'yuternaya optika [Computer Optics]*, 2025, vol. 49, no. 3, pp. 443–450. DOI: 10.18287/2412-6179-CO-1563, EDN JUAWIT (In Russian)
15. Fedorov V. A., Ogorodnikova O. M. Segmentatsiya ob"ektov tekhnicheskim zreniem v avtomatizirovannykh sistemakh upravleniya zheleznodorozhnym podvizhnym sostavom [Segmentation of Objects by Technical Vision in Automated Railway Rolling Stock Control Systems], *Avtomatika na transporte [Transport Automation Research]*, 2025, vol. 11, no. 3, pp. 239–249. DOI: 10.20295/2412-9186-2025-11-03-239-249. EDN IUNEIX (In Russian)
16. He K., et al. Deep Residual Learning for Image Recognition, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. DOI: 10.1109/CVPR.2016.90
17. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, *International Conference on Machine Learning (ICML)*, 2019. DOI: 10.48550/arXiv.1905.11946
18. Huang G., et al. Densely Connected Convolutional Networks, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017. DOI: 10.1109/CVPR.2017.243
19. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation, *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015. DOI: 10.1007/978-3-319-24574-4\_28
20. Zhou Z., et al. UNet++: A Nested U-Net Architecture for Medical Image Segmentation, *International Workshop on Deep Learning in Medical Image Analysis (DLMIA/MICCAI)*, 2018. DOI: 10.1007/978-3-030-00889-5\_1
21. Chen L.-C., et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, *European Conference on Computer Vision (ECCV)*, 2018. DOI: 10.48550/arXiv.1802.02611
22. Li R., et al. Multi-Attention-Network for Semantic Segmentation of Fine Resolution Remote Sensing Images, *IEEE Transactions on Geoscience and Remote Sensing (TGRS)*, 2022. DOI: 10.1109/TGRS.2021.3093977
23. Xie E., et al. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers, *Conference on Neural Information Processing Systems (NeurIPS)*, 2021. DOI: 10.48550/arXiv.2105.15203
24. Popov P. A., Ozerov A. V., Marshova A. S. Prospects of Autonomous Railway Transport Development, *BRICS Transport*, 2024, no. 3 (3), pp. 1–14. DOI: 10.46684/2024.3.4, EDN HPYXEZ