УДК 004.896:656.25

ПОМЕХОЗАЩИЩЕННОСТЬ ИНТЕЛЛЕКТУАЛЬНЫХ ТРАНСПОРТНЫХ СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ АНАЛИЗА ИЗОБРАЖЕНИЙ ЖЕЛЕЗНОДОРОЖНОЙ ИНФРАСТРУКТУРЫ

КУЛАГИН Максим Алексеевич, канд. техн. наук, доцент[1]; e-mail: maksimkulagin06@yandex.ru
ЯНЧЕНКО Герман Олегович, аспирант[1]; e-mail: ya@gyanchenko.ru
РОДИНА Дарья Михайловна, ведущий разработчик[2]; e-mail: dk.957@ya.ru
ПОЛЕГЕНЬКО Алексей Иванович, студент[1]; e-mail: screen.polegenko@mail.ru

[1]Кафедра «Управление и защита информации», Российский университет транспорта, Москва
[2]ПАО «ВТБ», Москва

В условиях активного внедрения интеллектуальных систем в железнодорожную отрасль возрастает значимость обеспечения их робастности к внешнему вмешательству, в том числе в форме скрытых атак на входные данные. Статья посвящена исследованию робастности современных архитектур нейронных сетей (*ResNet18*, *ResNet50*, *Vision Transformer* (*ViT*), сверточной нейронной сети и мультимодальной *GPT-4o*), применяемых для автоматического обнаружения дефектов на изображениях элементов железнодорожной инфраструктуры. Проведены эксперименты с генерацией скрытых возмущений с помощью универсального шума, созданного на базе ансамбля трансформеров. Рассмотрены две модификации атак (M1 и M*2*), позволяющие смоделировать реалистичные сценарии вмешательства в условиях ограниченного доступа к данным. Оценка качества моделей выполнялась как на «чистых» изображениях, так и в условиях добавленного шума. Результаты показывают, что, несмотря на высокую точность *ResNet50* на исходных данных, наибольшую робастность к возмущениям демонстрируют *ViT* и *GPT-4o*. Сделаны выводы о целесообразности выбора архитектуры не только по точности, но и по уровню робастности к шуму. Работа предлагает методику оценки робастности и практические рекомендации для разработки систем компьютерного зрения, предназначенных для эксплуатации в критически важных условиях железнодорожного транспорта.

Ключевые слова: интеллектуальные системы; железнодорожный транспорт; состязательные атаки; робастность; нейронные сети; компьютерное зрение; безопасность; мониторинг инфраструктуры.

DOI: 10.20295/2412-9186-2025-11-04-313-326

▼ Введение

В настоящее время интеллектуальные системы внедряются в различные процессы на железнодорожном транспорте, например для мониторинга состояния путей, управления движением и обеспечения безопасности [1–3].

В рамках данного исследования рассматриваются интеллектуальные транспортные системы (ИТС), которые используют техническое зрение. Требования к системам технического зрения (СТЗ) зависят от функций, которые эти системы выполняют, и задач, которые они решают.

В современных ИТС ключевую роль играют СТЗ в реализации беспилотного управления, особенно на железнодорожном транспорте. СТЗ используются для обнаружения препятствий на пути следования поезда, оценки обстановки и принятия решений в режиме реального времени без участия машиниста [4]. Критически важным требованием для таких беспилотных систем является способность видеть дальше и реагировать быстрее человека-машиниста [5]. При обнаружении опасного объекта современные системы могут не только предупредить оператора, но и автоматически задействовать тормоза без вмешательства человека. Все это делает техническое зрение базовым элементом обеспечения безопасности движения в будущих беспилотных поездах.

В условиях, когда полностью автономное управление еще не реализовано, СТЗ

используются для помощи машинисту и повышения уровня осведомленности о ситуации человека-оператора. Такие системы функционируют как электронный помощник машиниста: камеры, установленные на подвижном составе, непрерывно наблюдают путь, а нейронная сеть в бортовом вычислительном модуле анализирует видеопоток для поиска препятствий. Примером является бортовая система технического зрения, внедряемая на локомотивах РЖД [6, 7]. Она распознает на пути людей, животных, автомобили, другую технику, а также объекты железнодорожной инфраструктуры (сигналы светофоров, стрелки, тупиковые упоры и т. д.), которые могут представлять опасность для движения. Основная задача таких систем — предотвратить проезд на запрещающий сигнал, столкновения с преградами и иные инциденты, повышая безопасность движения на станциях и перегонах.

Отдельным направлением применения компьютерного зрения на транспорте являются задачи диагностики и мониторинга состояния инфраструктуры и подвижного состава. Здесь можно выделить системы, работающие в реальном времени, и системы, работающие в отсроченном режиме. В одном случае СТЗ интегрируется непосредственно в эксплуатационный процесс: обнаружение критических дефектов или отклонений сразу генерируют сигналы тревоги, позволяя немедленно принять меры. Например, с помощью панорамных видеокамер высокого разрешения выполняют автоматическое распознавание номеров вагонов, мониторинг температуры, визуализацию пантографа [8]. Такие СТЗ позволяют заблаговременно выявлять проблемы в состоянии подвижного состава до его прибытия в парк. С другой стороны, многие диагностические СТЗ действуют в режиме, близком к реальному времени, но все же предполагают анализ данных после их сбора. Например, система автоматического выявления дефектов рельс и стрелочных переводов, которая повышает безопасность и эффективность эксплуатации железных дорог [9]. В работе с использованием нейронных сетей удалось автоматически выявить изъяны железнодорожных путей, достигнув точности порядка 80–82 %.

В результате анализа источников литературы можно существующие СТЗ, применяемые на железнодорожном транспорте, классифицировать по нескольким ключевым признакам в зависимости от их назначения и условий работы:

1. По временному режиму функционирования:
- системы реального времени;
- системы, работающие в отложенном режиме.

2. По степени автоматизации и роли человека:
- полностью автономные;
- частично автоматизированные;
- информационно-мониторинговые.

3. По функциональному назначению:
- системы обеспечения безопасности движения включают в себя подсистемы обнаружения препятствий на пути, распознавания сигналов светофоров и знаков, а также системы обнаружения несанкционированного проникновения на пути;
- системы технической диагностики и обслуживания нацелены на обнаружение физических дефектов и отказов — сюда относятся комплексы мониторинга состояния железнодорожного пути, а также системы проверки подвижного состава во время движения поезда или на станциях;
- системы контроля состояния персонала и соблюдения регламента — например, камеры, следящие за бодрствованием и концентрацией машиниста (предупреждение засыпания, состояние опьянения) [10].

4. По месту установки и охвату обзора камер и датчиков:
- подвижной состав;
- элементы инфраструктуры;
- мобильные платформах (например, дроны, роботы или носимые устройства).

В критически важных задачах применения СТЗ возрастает внимание к вопросам их защищенности от потенциальных угроз. Незначительные визуальные модификации, незаметные для человека, могут привести к ошибочной классификации со стороны модели [11], что в условиях железнодорожного транспорта может иметь серьезные последствия для безопасности движения. Эти модификации могут

быть внесены преднамеренно в форме состязательных атак [12], что требует исследований в области оценки робастности моделей к подобным воздействиям. В статистическом смысле [13] робастность означает нечувствительность к малым отклонениям от предположений. В данной статье робастность модели определяется в контексте машинного обучения [11], что означает способность модели сохранять качество прогнозирования при воздействии возмущений на входные данные, включая шум и целенаправленные состязательные атаки. В данном исследовании под шумом понимается случайное, статистически описываемое отклонение яркости/цвета пикселей, не имеющее структуры.

СТЗ без специальных мер защиты уязвимы к скрытому вмешательству. В ответ на такую угрозу разрабатываются различные методы защиты моделей:

- фильтрация входных данных [14];
- детектирование аномалий [15];
- обучение с учетом возмущений (состязательное обучение) [16].

Одним из методов защиты моделей считается состязательное обучение, при котором в процесс тренировки включаются состязательные примеры. Реализация состязательного дообучения заметно повышает стабильность моделей искусственного интеллекта без потери точности. В работе [17] авторы показали, что модель классификации состояния инфраструктуры для высокоскоростных поездов, обученная на видеоданных, демонстрировала высокое качество работы при обработке незашумленных изображений, однако проявила сильное снижение качества (то есть оказалась уязвима к атакам) при добавлении к кадрам небольших изменений методом BIM [18]. BIM (Basic Iterative Method) — это итерационный вариант атаки FGSM (Fast Gradient Sign Method) [19], при котором к изображению многократно добавляют небольшие порции возмущения, вычисляемого по градиенту потерь модели, чтобы заставить модель ошибаться.

В работах [20, 21] рассматриваются состязательные атаки на системы машинного зрения и методы защиты от них. Анализируются уязвимости современных моделей ИИ, и обобщаются результаты исследований по повышению их робастности, включая состязательное обучение, фильтрацию и реконструкцию входных данных, а также архитектурные решения для снижения чувствительности к шуму. В [22] анализируются методы борьбы с шумом на изображениях в части обнаружения скрытой информации. Авторы предлагают улучшить метод комплексного стегоанализа путем учета качественных характеристик изображений (резкости, размытости, шума, контраста и энтропии) наряду с классическими стегоаналитическими методами.

Данные обзоры обобщают подходы по защите моделей ИИ от помех и злонамеренных воздействий, что актуально для железнодорожных интеллектуальных систем. На сегодняшний день не существует универсального решения, гарантирующего полную защиту нейронных сетей от всех видов атак. Каждая оборонительная мера имеет ограничения, и злоумышленники продолжают находить новые способы обхода защит. Таким образом, проблема обеспечения безопасности ИТС железнодорожного транспорта при скрытом внешнем воздействии остается актуальной и требует дальнейшего исследования.

В данной статье для оценки робастности моделей ИИ рассматривается задача определения дефектов в состоянии инфраструктуры по изображениям. В работе [23] предложена модифицированная глубокая сверточная нейронная сеть для автоматического распознавания и классификации дефектов рельсовых стыков по видеоизображениям, полученным вагономпутеизмерителем. Авторы применили метод переноса обучения и расширили обучающие данные с помощью искусственных аффинных преобразований изображений, чтобы повысить стабильность работы классификатора к различным положениям стыков. В статье [24] исследовано качество работы нейронной сети (YOLOv8), решающей задачу детектирования объектов железнодорожной инфраструктуры в реальном времени. Модель обучена на большом наборе данных (более 20 000 изображений) и протестирована на различных конфигурациях (разные разрешения входных изображений). Результаты показали высокую робастность алгоритма

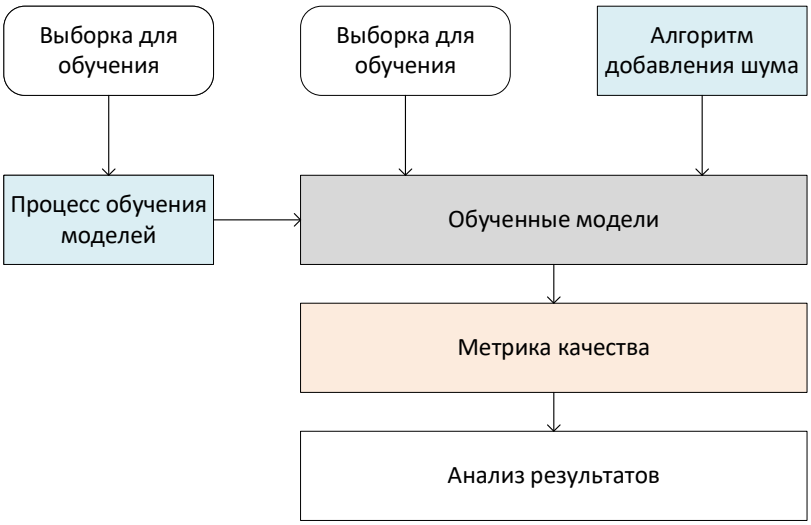


Рис. 1. Обобщенная схема эксперимента

YOLOv8 к изменениям условий среды — модель обнаруживает элементы пути при различном окружении и шуме. Использование специального аппаратного графического ускорителя существенно повысило скорость обработки, что имеет значение для мониторинга инфраструктуры в режиме реального времени.

Целью настоящей работы является экспериментальное исследование робастности различных нейросетевых архитектур, используемых для автоматического обнаружения дефектов железнодорожной инфраструктуры, в условиях скрытого вмешательства, моделируемого с помощью «зашумления» изображений.

Решаются следующие задачи:

- создание и обучение моделей для обнаружения дефектов рельсов по изображению;
- создание алгоритмов добавления помех на изображения;
- оценка качества разработанных моделей на «чистых» и «зашумленных» изображениях;
- анализ архитектур моделей в условиях помех.

Схема эксперимента иллюстрирует процесс оценки робастности моделей ИИ к шуму (рис. 1).

1. Используемые модели обнаружения дефектов рельсов по изображению

Даны обучающий $D^{train} = \left\{\left(x_i, y_i\right)\right\}_{i=1}^{N}$ и тестовый $D^{test} = \left\{\left(x_j^{test}, y_j^{test}\right)\right\}_{j=1}^{M}$ наборы данных, в которых $x_i, x_j^{test} \in R^{H\times W\times 3}$ — изображения участка рельсового пути, где $H \times W \times 3$ — это высота, ширина и количество каналов в RGB-изображении, $y_i, y_j^{test} \in \{0,1\}$ — метка, N, M — количество изображений в обучающем и тестовом наборе данных соответственно. Количество изображений с дефектами ($y_i = 1$) и без дефектов ($y_i = 0$) одинаковое (рис. 2).

Требуется построить параметризованную модель $f_и: R^{HxWx3} \to \{0,1\}$, которая аппроксимирует апостериорное распределение $P(y=1|x)$. Подбор параметров модели производится с использованием алгоритма градиентного спуска [26]. Суть данного алгоритма заключается в итеративном обновлении параметров θ в направлении, противоположном градиенту $L(\theta)$ функции потерь. На каждой итерации вычисляется градиент функции потерь по параметрам модели $\nabla_\theta L(\theta)$, после чего параметры обновляются по правилу:

$$\theta_{t+1} = \theta_t + \eta \nabla_\theta L(\theta_t), \quad (1)$$

где t — итерация обучения модели;

$\eta > 0$ — скорость обучения, определяющая величину шага.

Таким образом, модель последовательно уточняет свои параметры, стремясь минимизировать выбранную функцию ошибки. В качестве оптимизируемой функции ошибки используется бинарная кросс-энтропия:

$$L(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(p_i\right) + \left(1-y_i\right)\log\left(1-p_i\right)\right], \quad (2)$$

Рис. 2. Примеры изображений из набора данных
(в верхнем ряду с дефектами, а в нижнем — без дефектов). Набор данных взят из [25]

где p_i — прогноз обученной модели. После решения задачи оптимизации получаются параметры $\hat{\theta}$ и обученная модель $\hat{f}(x) = f_{\hat{\theta}}(x)$.

В рамках исследования были задействованы пять различных архитектур нейронных сетей, охватывающих как сверточные, так и модели трансформеров. Это позволило провести сравнительный анализ робастности моделей различных классов к скрытым помехам.

1. Модель семейства *Vision Transformer*, *ViT-B/16*, разбивающая входное изображение на фрагменты размером 16 × 16 пикселей и обрабатывающая их с помощью механизма внимания. Обеспечивает баланс между вычислительной сложностью и качеством классификации.

2. Классические сверточные архитектуры с остаточными связями [27]. *ResNet18* представляет собой компактную модель, состоящую из 18 слоев, а *ResNet50* — более глубокую, с 50 слоями, обеспечивающую высокую точность при наличии достаточного количества обучающих данных. Обе модели использовались с предварительно обученными весами (*ImageNet*) и дообучались на целевом датасете.

3. Мультимодальная модель *GPT-4o*, способная обрабатывать как текстовые, так и визуальные входные данные. В рамках эксперимента использовалась для классификации изображений с использованием prompt-инжиниринга, без дополнительного дообучения модели.

4. Сверточная нейронная сеть (*CNN*). Архитектура предлагаемой нейронной сети представляет собой *CNN*, ориентированную на задачу бинарной классификации изображений. Модель состоит из четырех сверточных блоков, каждый из которых включает в себя последовательность из свертки с ядром 3 × 3, нормализации по батчу (*BatchNorm2d*) и функции активации *ReLU*, за которой следует операция подвыборки (*MaxPooling2d*) для первых трех блоков и адаптивное усреднение (*AdaptiveAvgPool2d*) в последнем блоке. Последовательно увеличивается количество фильтров от 32 до 256, что позволяет эффективно извлекать признаки различной степени абстракции. После экстракции признаков из изображения выходное тензорное представление проходит через слой выравнивания (*Flatten*) и линейный полносвязный слой, преобразующий 256 признаков в скалярное значение, на которое накладывается сигмоид-функция. Это позволяет интерпретировать выход как вероятность принадлежности к положительному классу. Архитектура сбалансирована по глубине и числу параметров, обеспечивая хорошее соотношение между вычислительной эффективностью и качеством классификации.

2. Алгоритм добавления помех на изображения

В D^{test} для каждого x_j^{test} создается модифицированная версия $\tilde{x}_j$:

$$\tilde{x}_j = x_j^{test} + \delta_j, \quad (3)$$

где δ_j — вектор шума, который создается по алгоритму, рассмотренному ниже.

Основная задача в рамках данного исследования заключается в оценке робастности нейронных сетей различной архитектуры, решающих задачу бинарной классификации. Робастность рассчитывается по следующей формуле:

$$R_p = 1 - \frac{\|\Delta Q\|_p}{\|Q^{clean}\|_p}, \quad (4)$$

где Q^{clean} — набор метрик качества на тестовых данных;

$\Delta Q = Q^{clean} - Q^{noisy}$ — показывающий ухудшение модели;

p — векторная норма, в рамках данного исследования была выбрана Евклидова норма $p = 2$. В данном случае чем ближе R_p к 1, тем стабильнее работает модель при наличии возмущений в данных.

В рамках настоящего исследования особое внимание уделяется анализу робастности моделей классификации к скрытому вмешательству в исходные данные. Для имитации возможных атак на систему мониторинга рельсового пути, в частности — скрытых модификаций изображений, используется метод генерации состязательных примеров.

Суть метода заключается в последовательной адаптации входного изображения таким образом, чтобы минимальные и визуально неразличимые изменения вызывали существенные сдвиги в выходных предсказаниях модели. Генерация осуществляется с использованием ансамбля суррогатных моделей и стохастических преобразований, что позволяет добиться переносимости атаки на разные архитектуры классификаторов.

В отличие от классических атак типа *FGSM*, используемый подход формирует возмущение, приближая выходное изображение к целевому изображению с использованием процесса добавления шума (рис. 3). В качестве функции расчет расстояния между двумя векторами, сформированными на выходе ансамбля моделей — текущего «зашумленного» $\tilde{x}_k^{sou}$ и целевого $\tilde{x}_k^{tar}$ изображений, выступает косинусное расстояние $\mathcal{L}\left(f\left(\tilde{x}_k^{sou}\right), f\left(\tilde{x}_k^{tar}\right)\right)$. Косинусное расстояние для сравнения векторов эмбеддингов (скрытых признаков) исходного и зашумленного изображений рассчитывается по формуле:

$$d(u,v) = 1 - \frac{u \cdot v}{\|u\|_2 \|v\|_2}, \quad (5)$$

где $u \cdot v$ — скалярное произведение векторов локальных и глобальных эмбеддингов;

$\|u\|_2 \|v\|_2$ — евклидовы нормы векторов.

Ансамбль моделей в исследовании использует три предварительно обученные архитектуры из семейства *Vision Transformer* (*ViT*) [25] ($m = 3$), каждая из которых реализована с различными конфигурациями параметров:

- ViT-B/16 — базовая модель, использующая разбиение входного изображения на непересекающиеся фрагменты размером 16 × 16 пикселей (patch embeddings), что обеспечивает баланс между вычислительной эффективностью и качеством извлечения признаков;
- ViT-B/32 — облегченная версия с размером фрагмента 32 × 32 пикселя, обладающая пониженной чувствительностью к локальным шумам за счет более грубого представления;
- ViT-g-14 — крупномасштабная архитектура, использующая фрагменты размером 14 × 14 пикселей и высокоразмерные векторы признаков, что способствует повышенной обобщающей способности и точности при обработке сложных входных данных.

Алгоритм 1. Алгоритм добавления шума

Вход: $x^{clean}, x^{tar}, n, \alpha, \varepsilon, m$

Выход: x^{noisy}

$x_0^{sou} \leftarrow x^{clean}; \delta_0 \leftarrow 0$

for $k \leftarrow 0$ *to* $n-1$ *do*

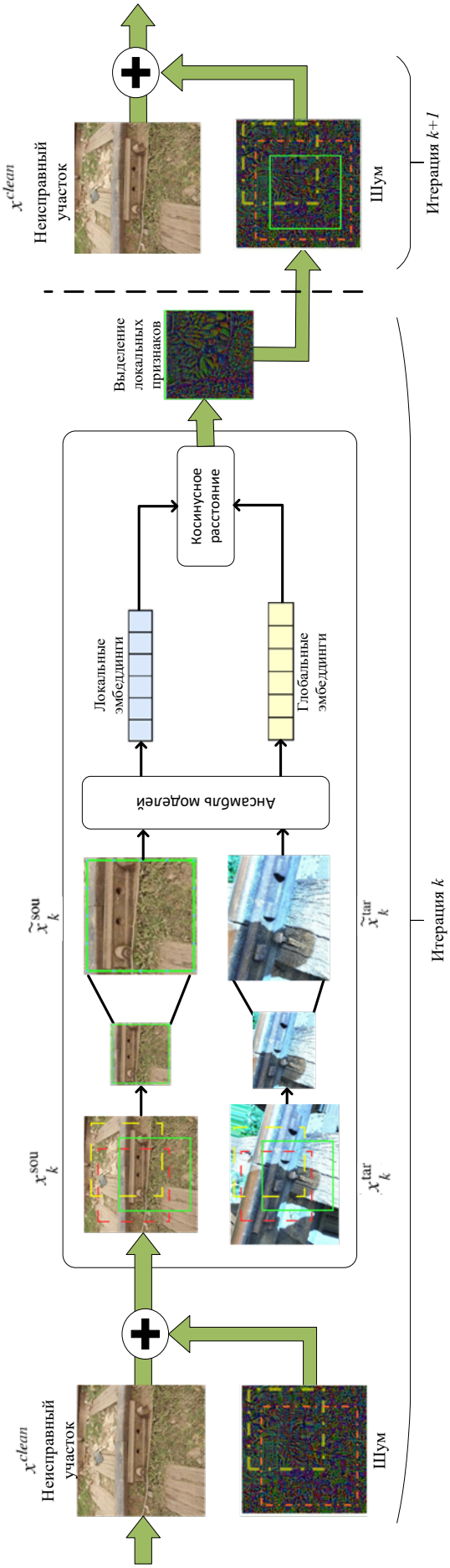


Рис. 3. Процесс добавления шума на изображения

$$\tilde{x}_k^{sou} \leftarrow \mathcal{F}_s(x_k^{sou}), \quad \tilde{x}_k^{tar} \leftarrow \mathcal{F}_s(x_k^{tar});$$

$$L \leftarrow \frac{1}{m}\sum_{j=1}^{m} L\left(f(\tilde{x}_k^{sou}), f(\tilde{x}_k^{tar})\right);$$

$$g_k \leftarrow \frac{1}{m}\nabla_{\tilde{x}_k^{sou}} \sum_{j=1}^{m} L\left(f(\tilde{x}_k^{sou}), f(\tilde{x}_k^{tar})\right);$$

$$\delta_{k+1} \leftarrow Clip(\delta_k + \alpha \cdot \text{sign}(g_k), -\varepsilon, \varepsilon);$$

$$x_k^{sou} \leftarrow \tilde{x}_k^{sou} + \delta_{k+1};$$

end for

$$x^{noisy} \leftarrow x_n^{sou},$$

где x^{clean} — исходное «чистое» изображение неисправного участка пути (оригинал без вмешательств);

x_k^{sou} — «зашумленное» на k-й итерации изображение;

x_k^{tar} — изображение, к которому подбирается шум, чтобы минимизировать различия признаков на k-й итерации;

$\tilde{x}_k^{sou}$ — фрагменты «зашумления» на k-й итерации изображение;

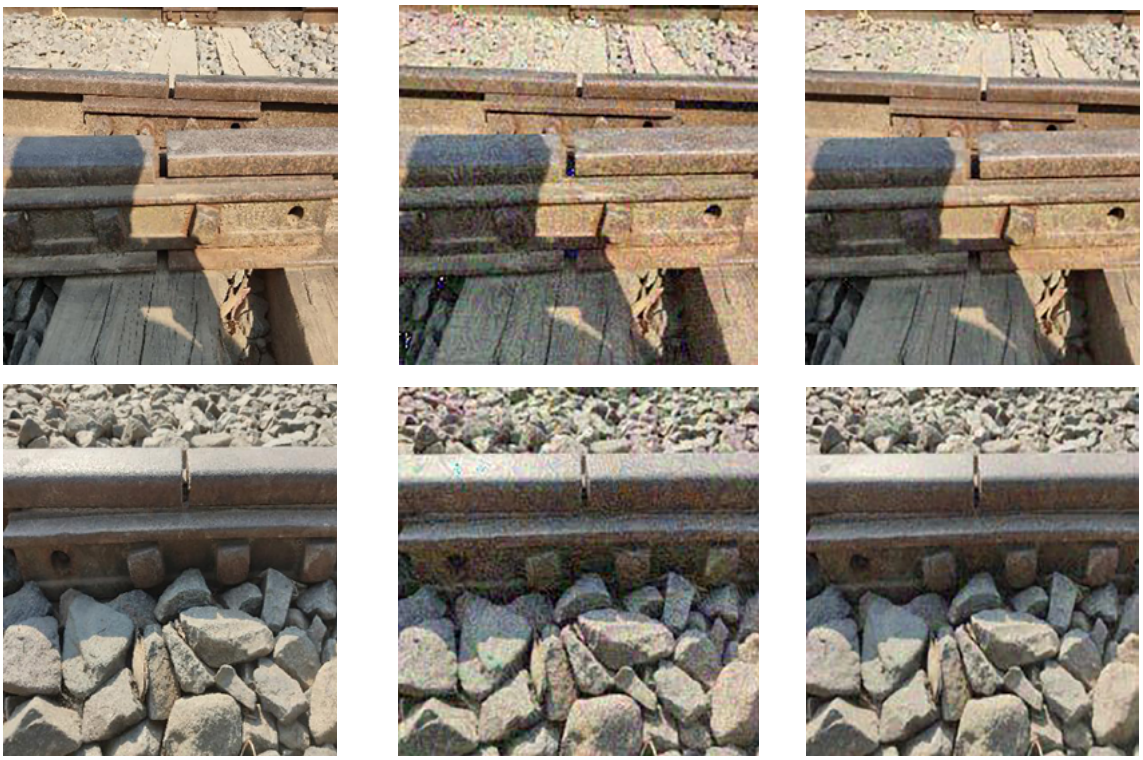
$\tilde{x}_k^{tar}$ — фрагменты целевого изображение на k-й итерации.

В данном исследовании рассматривается две модификации Алгоритма-1.

Модификация-1 (M1) заключается в создании одного универсального шума δ_k, который вычисляется с использованием градиентов по всем изображениям. Смысл M1 заключается в том, что для каждого изображения с классом $y_i = 1$ добавляется один универсальный шум, который вычисляется по аналогии с шагом 2 в A1, для всех примеров, в которых $y_i = 0$.

Модификация-2 (M2) заключается в создании одного универсального шума δ_k, который вычисляется как среднее значение шума, вычисленного по всем изображениям. Для каждого изображения с классом $y_i = 1$ добавляется один универсальный шум, который вычисляется по аналогии с шагом 2 в A1, для каждого примера в отдельности с $y_i = 0$. Затем шум усредняется по всем примерам.

В M1 и M2 возмущение создается однократно и используется многократно, что делает его приближенным к реальным сценариям атак, когда злоумышленник имеет ограниченный доступ к данным (рис. 4).



а *б* *в*

Рис. 4. Результат работы алгоритма добавления шума:
а — исходное изображение; *б* — шум, наложенный с использованием алгоритма и модификацией M1;
в — шум, наложенный с использованием алгоритма и модификацией M2. Набор данных из [25]

3. Оценка качества разработанных моделей на «чистых» и на «зашумленных» изображениях

Оценка качества модели производится по группе метрик $Q \in R^6$. Каждый элемент Q — это одна метрика, которая лежит в диапазоне от 0 до 1. В рамках данного исследования использовался следующий набор из 6 метрик для оценки бинарного классификатор [28]: *Accuracy*, *Precision*, *Recall*, *F1-score*, *ROC AUC*, *PR AUC*.

Accuracy — доля правильно классифицированных объектов среди всех;

Precision — доля верных положительных предсказаний среди всех положительных предсказаний модели;

Recall — доля верно найденных положительных случаев среди всех реально положительных случаев;

F1-score — сбалансированная оценка, объединяющая *Precision* и *Recall*;

ROC AUC — способность модели различать положительный и отрицательный классы, основываясь на площади под ROC-кривой;

PR AUC — это показатель качества модели, определяемый по площади под кривой Precision-Recall; он особенно важен при работе с несбалансированными данными.

Оценка качества производилась по обученной модели $\hat{f}$, которая предсказывала вероятность события $y_i = 1$, а затем бинаризировалась $\hat{y}_j = I[\hat{f}(x_j \geq \tau]$, где τ — порог бинаризации.

Для оценки базовой точности классификации в условиях отсутствия помех был проведен эксперимент, в котором все модели тестировались на исходных изображениях без наложения скрытых возмущений. Результаты сравнения представлены на рис. 5. Из диаграмм видно, что на «чистых» данных наибольшую точность (*Accuracy*) и обобщающую способность (*F1-score*, *ROC AUC*, *PR AUC*) демонстрирует модель *ResNet*-50, за ней следуют *ResNet*-18 и кастомная *CNN*. Модель *ViT-B/16* также показала стабильные показатели, особенно по метрике *Precision*, но уступила *ResNet*-50 по *F1-score* и *Recall*.

Эти результаты указывают на то, что глубокие сверточные архитектуры с предобученными весами (такие как *ResNet*-50) способны эффективно классифицировать дефекты железнодорожных креплений при наличии ограниченного обучающего набора. Полученные значения метрик служат отправной точкой для последующего анализа робастности моделей к скрытым вмешательствам, представленного на рис. 6.

Проведем сравнение эффективности различных моделей при воздействии скрытых

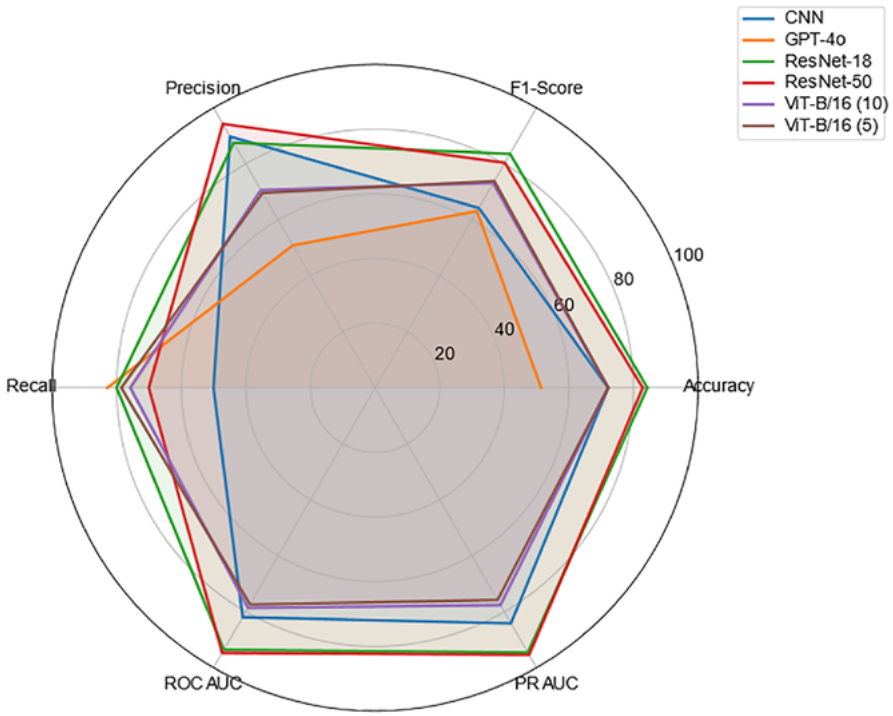
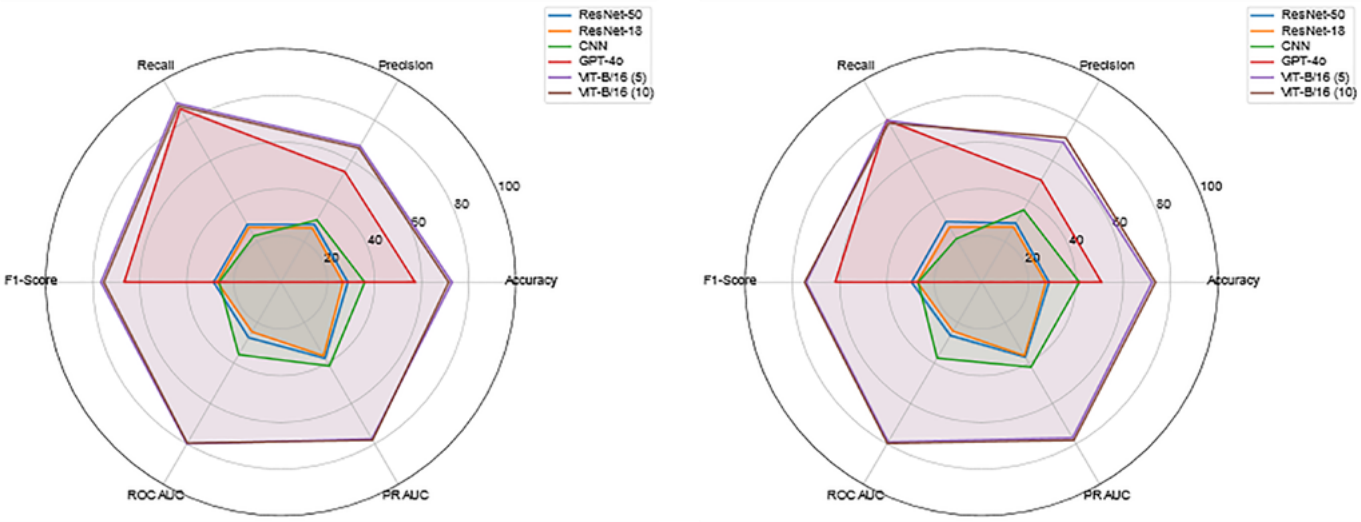


Рис. 5. Результаты работы моделей на «чистых» данных



а *б*

Рис. 6. Результаты работы моделей при наличии шума для сценариев: *а* — M1; *б* — M2

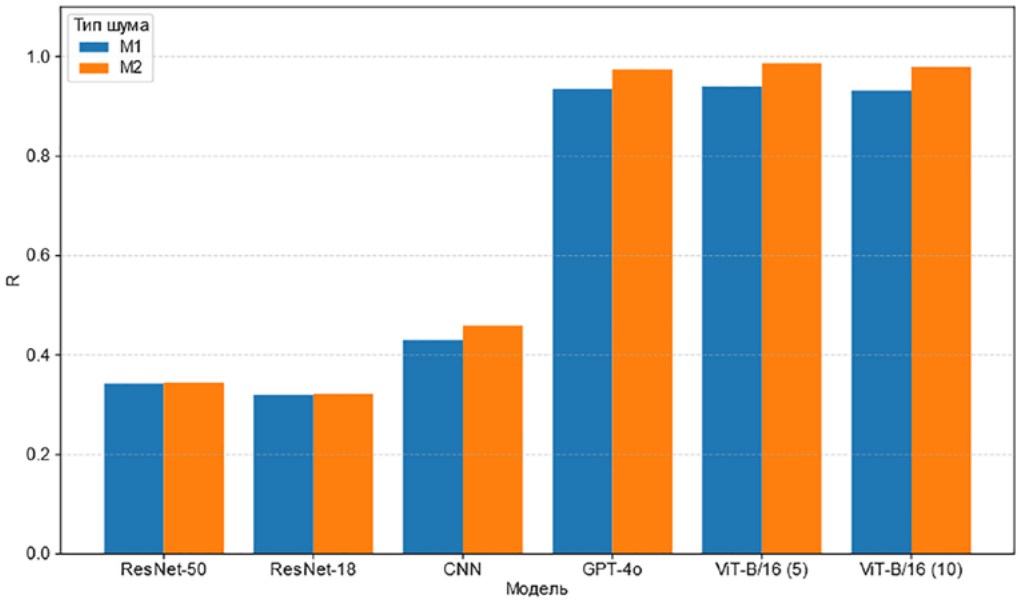


Рис. 7. Сравнение робастности моделей

возмущений, сгенерированных по двум сценариям: M1 и M2. Результаты показывают, что добавление шума существенно снижает значения всех ключевых метрик. Наименее чувствительными оказались модели *ViT-B*/16 и *GPT-4o*. Таким образом, можно сделать вывод, что выбор архитектуры существенно влияет на робастность модели. Применение *ResNet*-архитектур с предварительным обучением делает их менее подверженными внешним воздействиям, что особенно важно для критических задач транспортной инфраструктуры. Отсутствие значения метрик *ROC AUC* и *PR AUC* для *GPT-4o* связано с тем, что *GPT-4o* выдает на выходе 0 или 1.

4. Анализ архитектур моделей на предмет робастности в условиях помех

На графике (рис. 7) представлено сравнение робастности различных моделей (*ResNet18*, *ResNet50*, *GPT-4o, ViT* и *CNN*). Чем выше значение *R*, тем менее чувствительна модель к шуму во входных данных.

Как видно, модель *ViT* и *GPT-4o* демонстрирует высокую робастность, тогда как *CNN*, *ResNet18* и *ResNet50* показывают более заметное снижение качества при наличии шума. Такие результаты подтверждают потенциал *Vision Transformer*-архитектур в задачах, требующих надежности при «зашумлении» входных изображений.

Проведенный анализ показал, что:

- на «чистых» данных наилучшую точность демонстрируют глубокие сверточные модели ResNet-50;
- при добавлении шума модели на основе *Vision Transformer* (*ViT-B/16*) и *GPT-4o* проявляют наибольшую робастность к шуму;
- предложенные модификации универсального шума M1 и M2 позволяют эффективно моделировать реалистичные атаки без необходимости доступа к конкретным данным модели;
- использование архитектур, демонстрирующих высокую робастность, критически важно при разработке ИИ-систем, применяемых в условиях эксплуатации железнодорожного транспорта.

Заключение

В условиях активного внедрения интеллектуальных систем в железнодорожный транспорт проблема их робастности к внешнему скрытому вмешательству становится все более актуальной. Современные методы СТЗ, основанные на глубоком обучении, показывают высокие результаты при решении задач мониторинга инфраструктуры, однако их чувствительность к состязательным атакам ставит под угрозу надежность работы в реальных, зачастую сложных условиях эксплуатации.

Настоящее исследование продемонстрировало, что архитектура модели играет ключевую роль в обеспечении робастности интеллектуальных систем к скрытым воздействиям. Несмотря на то, что модели семейства *ResNet*, в особенности *ResNet50*, достигли наивысших значений точности на «чистых» данных, их робастность к шуму оказалась существенно ниже по сравнению с моделями на основе *Vision Transformer* (*ViT-B/16*) и мультимодальной *GPT-4o*. Это позволяет утверждать, что при проектировании систем для задач мониторинга, где возможны попытки внешнего вмешательства, следует отдавать предпочтение архитектурам, демонстрирующим робастность даже в ущерб незначительному снижению точности на чистых данных.

Эксперименты с добавлением универсального шума по разработанным модификациям (M1 и M2) показали, что даже слабые визуальные возмущения могут существенно снизить качество классификации у большинства моделей. Это подчеркивает необходимость интеграции механизмов защиты уже на этапе проектирования интеллектуальных систем. Модификация M2, основанная на усреднении индивидуальных возмущений, оказалась особенно эффективной для имитации реальных сценариев атак, когда злоумышленник имеет ограниченный доступ к системе, но обладает знаниями о ее архитектуре и типичных данных.

Важно отметить, что мультимодальная модель *GPT-4o*, несмотря на отсутствие дообучения на конкретном датасете, продемонстрировала высокую робастность. Это открывает новые перспективы для применения подобных моделей в условиях недостатка обучающих данных или необходимости адаптации ИИ-систем под разные задачи без переобучения.

Таким образом, результаты работы подчеркивают необходимость пересмотра приоритетов при разработке ИИ-систем для критически важных сфер, таких как железнодорожный транспорт. Не только точность модели, но и ее робастность к внешним воздействиям должны становиться определяющим критерием при выборе архитектурных решений и технологий обучения.

Важным направлением будущих исследований является разработка активных стратегий противодействия атакам, что включает:

- внедрение методов предварительной фильтрации и нормализации входных изображений;
- создание специализированных моделей-детекторов, отслеживающих аномалии в потоках данных.

Не менее значимым направлением остается повышение вычислительной эффективности и адаптация защищенных моделей для работы в режиме реального времени. Особенно акту-

ально это в системах, где задержка в несколько секунд может повлечь за собой серьезные последствия, например при выявлении дефектов на ходу поезда или в условиях круглосуточного мониторинга.

Таким образом, настоящее исследование закладывает основу для дальнейшего совершенствования ИИ-инструментов, робастных к внешнему вмешательству, и служит практическим ориентиром при выборе моделей для задач обеспечения безопасности в железнодорожной отрасли. Полученные результаты могут быть применены при разработке новых стандартов и требований к интеллектуальным системам мониторинга, а также в образовательных программах по кибербезопасности, машинному обучению и транспортной инженерии.

Благодарности

Работа выполнена за счет бюджетного финансирования в рамках государственного задания от 20.03.2025 № 103-00001-25-02. ▲

Список источников

1. Кулагин М. А. Обобщение опыта решения задач предиктивной аналитики на железнодорожном транспорте / М. А. Кулагин, В. Г. Сидоренко // Наука и техника транспорта. — 2024. — № 4. — С. 55–62.
2. Сидоренко В. Г. Интеллектуальная система обнаружения нарушений в соблюдении требований безопасности при работах на объектах железнодорожной инфраструктуры / В. Г. Сидоренко, М. А. Кулагин, Д. М. Родина // Автоматика на транспорте. — 2025. — Т. 11. — № 1. — С. 55–65. — DOI: 10.20295/2412-9186-2025-11-01-55-65.
3. Малинский С. В. Автоматическое определение границ опасных участков железнодорожного пути / С. В. Малинский, А. В. Абрамов, В. О. Шарова // Интеллектуальные транспортные системы: материалы IV Международной научно-практической конференции, Москва, 22 мая 2025 года. — Москва: Российский университет транспорта (МИИТ), 2025. — С. 624–630. — DOI: 10.30932/9785002587582-2025-624-630.
4. Ашрафзянов А. М. Обнаружение и распознавание препятствий перед автомобилем на основе обработки видеоизображений / А. М. Ашрафзянов, М. П. Шлеймович // Вестник Казанского государственного технического университета им. А.Н. Туполева. — 2014. — № 2. — С. 197–202.
5. Баранов Л. А. Методология обоснования требований безопасности при использовании систем технического зрения в интеллектуальных системах управления движением поездов / Л. А. Баранов, П. Ф. Бестемьянов, Е. П. Балакина, А. Л. Охотников // Интеллектуальные транспортные системы: материалы Международной научно-практической конференции, Москва, 26 мая 2022 года. — М.: Российский университет транспорта, 2022. — С. 54–58.
6. Охотников А. Л. Проекты систем технического зрения для автоматического управления движением / А. Л. Охотников // Автоматика, связь, информатика. — 2023. — № 3. — С. 21–24. — DOI: 10.34649/AT.2023.3.3.003.
7. Озеров А. В. Техническое зрение в составе систем автоматического управления движением поездов / А. В. Озеров, А. С. Маршова // Перспективные информационные технологии (ПИТ 2022): труды Международной научно-технической конференции, Самара, 18–21 апреля 2022 года / под ред. С. А. Прохорова. — Самара: Издательство Самарского научного центра РАН, 2022. — С. 201–205.
8. Lisanti G. A Multi-Camera Image Processing and Visualization System for Train Safety Assessment / G. Lisanti, S. Karaman, D. Pezzatini // arXiv preprint. — 2015. — arXiv:1507.07815. — DOI: 10.48550/arXiv.1507.07815.
9. Saritas M. M. et al. Railway Track Fault Detection with ResNet Deep Learning Models // 2023 International Conference on Intelligent Systems and New Applications (ICISNA'23). — 2023.
10. Сунь Х. Обзор современных систем технического зрения, применяемых в транспортной отрасли / Х. Сунь, С. Чжуан, А. А. Костров // Современные наукоемкие технологии. — 2024. — № 9. — С. 69–73. — DOI: 10.17513/snt.40150.
11. Goodfellow I. Explaining and Harnessing Adversarial Examples / I. Goodfellow, J. Shlens, C. Szegedy // arXiv preprint. — 2015 (submitted 20 Dec 2014, revised 20 Mar 2015). — arXiv:1412.6572. — DOI: 10.48550/arXiv.1412.6572.
12. Григоренко А. Г. Обзор методов защиты от адверсиальной атаки One Pixel в системах машинного обучения / А. Г. Григоренко, Н. А. Васильев, Д. С. Ситдиков // Системы интеллектуального управления и искусственный интеллект: теория и практика: сборник трудов II национальной научно-практической конференции, Санкт-Петербург, 27 июня 2024 года. — СПб.: Федеральное государственное бюджетное образовательное учреждение высшего образования Государственный университет морского и речного флота им. адмирала С. О. Макарова, 2024. — С. 30–36.
13. Хьюбер П. Робастная статистика / П. Хьюбер. М.: Мир,1984.
14. Goodfellow I. J. Making machine learning robust against adversarial inputs / I. J. Goodfellow, J. Shlens, C. Szegedy // Communications of the ACM. — 2018. — Vol. 61. — Iss. 7. — Pp. 56–66. — DOI: 10.1145/3134599.
15. Василенко, М. С. Алгоритм машинного обучения для детектирования выбросов и аномалий / М. С. Василенко, А. С. Копырин // Modeling of Artificial Intelligence. — 2019. — № 6-1. — С. 13–18. — DOI: 10.13187/mai.2019.1.13.
16. Легашев Л. В. Методика построения устойчивой системы защиты на основе состязательного машинного обучения в беспроводных сетях 6G / Л. В. Легашев, Л. С. Гришина // Вопросы кибербезопасности. — 2023. — № 2(54). — С. 99–108. — DOI: 10.21681/2311-3456-2023-2-99-108.

17. Hou X. High-Speed Rail Operating Environment Recognition Based on Neural Network and Adversarial Training / X. Hou et al. //2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). — IEEE, 2019. — Pp. 840–847. — DOI: 10.1109/ICTAI.2019.00120.
18. Голдобин И. А. Влияние шумов на алгоритмы цифровой обработки изображений / И. А. Голдобин, Е. И. Климова // Актуальные вопросы развития современной цифровой среды: сборник статей по материалам научно-технической конференции молодых ученых, Москва, 14–16 апреля 2021 года. — Волгоград: Сириус, 2021. — С. 396–402.
19. Котенко И. В. Атаки и методы защиты в системах машинного обучения: анализ современных исследований / И. В. Котенко, И. Б. Саенко, О. С. Лаута и др. // Вопросы кибербезопасности. — 2024. — № 1(59). — С. 24–37. — DOI: 10.21681/2311-2024-1-24-37.
20. Костюмов В. В. Обзор и систематизация атак уклонением на модели компьютерного зрения / В. В. Костюмов // International Journal of Open Information Technologies. — 2022. — Т. 10. — № 10. — С. 11–20.
21. Potapov A. K. Vulnerabilities of Artificial Intelligence Systems / A. K. Potapov, V. G. Sidorenko // 2024 International Conference "Quality Management, Transport and Information Security, Information Technologies" (QM&TIS&IT). — IEEE, 2024. — Pp. 84–87. — DOI: 10.1109/QMTISIT63393.2024.10762915.
22. Грачев Я. Л. Использование качественных характеристик изображения для комплексного стегоанализа / Я. Л. Грачев, В. Г. Сидоренко // Надежность. — 2025. — Т. 25. — № 1. — С. 67–74. — DOI: 10.21683/1729-2646-2025-25-1-67-74.
23. Орлов С. П. Глубокая нейронная сеть для диагностики элементов железнодорожного рельсового пути / С. П. Орлов, Н. А. Ефимушкин, Н. В. Ефимушкина // Вестник Самарского государственного технического университета. Серия: Технические науки. — 2022. — Т. 30. — № 1(73). — С. 63–74. — DOI: 10.14498/tech.2022.1.4.
24. Федоров В. А. Обнаружение железнодорожной инфраструктуры на основе YOLOv8 с использованием нейронного процессора / В. А. Федоров // Доклады Российской академии наук. Математика, информатика, процессы управления. — 2024. — Т. 520. — № S2. — С. 49–56. — DOI: 10.31857/S2686954324700371.
25. Eunus S. I. ECARRNet: An Efficient LSTM-Based Ensembled Deep Neural Network Architecture for Railway Fault Detection / S. I. Eunus, S. Hossain, A. E. M. Ridwan, A. Adnan et al. // AI. — 2024. — Vol. 5. — Iss. 2. — Pp. 482–503. — DOI: 10.3390/ai5020024.
26. Goodfellow I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville. — Cambridge, MA: MIT Press, 2016. — DOI: 10.1007/s10710-017-9314-z.
27. Han K. A Survey on Visual Transformer / K. Han, Y. Wang, H. Chen // arXiv preprint. — 2020. — DOI: 10.48550/arXiv.2012.12556.
28. Powers D. M. W. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation / D. M. W. Powers // Journal of Machine Learning Technologies. — 2011. — Vol. 2. — Iss. 1. — Pp. 37–63. — DOI: 10.48550/arXiv.2010.16061.

TRANSPORT AUTOMATION RESEARCH, 2025, Vol. 11, No. 4, pp. 313–326
DOI: 10.20295/2412-9186-2025-11-04-313-326

Robustness of Intelligent Transport Systems Employing Computer Vision for the Analysis of Railway Infrastructure Images

Information about authors
Kulagin M. A., PhD in Engineering, Associate Professor[1].
E-mail: maksimkulagin06@yandex.ru
Yanchenko G. O., Postgraduate Student[1]. E-mail: ya@gyanchenko.ru
Rodina D. M., Senior Developer[2]. E-mail: dk.957@ya.ru
Polegenko A. I., Student[1]. E-mail: screen.polegenko@mail.ru

[1]"Control and Information Security" Department, Russian University of Transport, Moscow
[2]TI "VTB", Moscow

Abstract: With the growing integration of AI systems in the railway sector, it is becoming essential to ensure their robustness against external interference, particularly subtle alterations in input data. This paper analyses the resilience of various neural network architectures, including ResNet18, ResNet50, Vision Transformer, a convolutional neural network, and the GPT-4o multimodal model, which are used for automated fault detection in the analysis of railway infrastructure images. Experiments have been conducted using adversarial disturbances generated via universal noise derived from a set of transformers. Two attack modifications were employed to simulate real-world interference scenarios with limited data availability. The performance of the models has been evaluated on both untainted images and those compromised by overlaid noise. The outcomes indicate that while ResNet50attains maximum accuracy on unmodified data, ViT and GPT-4o demonstrate greater resilience to adversarial disturbances. The research emphasizes the importance of selecting model architectures based on both their accuracy and their robustness against distortions. This paper introduces a method for assessing robustness and offers practical recommendations for developing AI systems designed for application in railway environments where safety is critical.

Keywords: AI systems; railway transport; adversarial attacks; robustness; neural networks; computer image; safety; infrastructure monitoring.

References

1. Kulagin M. A., Sidorenko V. G. Obobshchenie opyta resheniya zadach prediktivnoy analitiki na zheleznodorozhnom transporte [Generalization of experience in solving predictive analytics problems in railway transport]. *Nauka i tekhnika transporta* [Science and Technology of Transport]. 2024, Iss. 4, pp. 55–62. (In Russian)
2. Sidorenko V. G., Kulagin M. A., Rodina D. M. Intellektual'naya sistema obnaruzheniya narusheniy v soblyudenii trebovaniy bezopasnosti pri rabotakh na ob"ektakh zheleznodorozhnoy infrastruktury [Intelligent system for detecting violations of safety requirements during work on railway infrastructure objects]. *Avtomatika na transporte* [Automation on Transport]. 2025, vol. 11, Iss. 1, pp. 55–65. DOI: 10.20295/2412-9186-2025-11-01-55-65. (In Russian)
3. Malinskiy S. V., Abramov A. V., Sharova V. O. Avtomaticheskoe opredelenie granits opasnykh uchastkov zheleznodorozhnogo puti [Automatic determination of boundaries of hazardous sections of railway track]. *Intellektual'nye transportnye sistemy: materialy IV Mezhdunarodnoy nauchno-prakticheskoy konferentsii, Moskva, 22 maya 2025 goda* [Intelligent Transport Systems: Proc. IV Int. Sci.-Pract. Conf., Moscow, 22 May 2025]. Moscow: Rossiyskiy universitet transporta (MIIT) Publ., 2025, pp. 624–630. DOI: 10.30932/9785002587582-2025-624-630. (In Russian)
4. Ashrafzyanov A. M., Shleymovich M. P. Obnaruzhenie i raspoznavanie prepyatstviy pered avtomobilem na osnove obrabotki videoizobrazheniy [Detection and recog-

nition of obstacles in front of a vehicle based on video image processing]. *Vestnik Kazanskogo gosudarstvennogo tekhnicheskogo universiteta im. A. N. Tupoleva* [Bulletin of Kazan State Technical University named after A. N. Tupolev]. 2014, Iss. 2, pp. 197–202. (In Russian)

5. Baranov L. A., Bestem'yanov P. F., Balakina E. P., Okhotnikov A. L. Metodologiya obosnovaniya trebovaniy bezopasnosti pri ispol'zovanii sistem tekhnicheskogo zreniya v intellektual'nykh sistemakh upravleniya dvizheniem poezdov [Methodology for justifying safety requirements when using computer vision systems in intelligent train traffic control systems]. *Intellektual'nye transportnye sistemy: materialy Mezhdunarodnoy nauchno-prakticheskoy konferentsii, Moskva, 26 maya 2022 goda* [Intelligent Transport Systems: Proc. Int. Sci.-Pract. Conf., Moscow, 26 May 2022]. Moscow: Rossiyskiy universitet transporta Publ., 2022, pp. 54–58. (In Russian)
6. Okhotnikov A. L. Proekty sistem tekhnicheskogo zreniya dlya avtomaticheskogo upravleniya dvizheniem [Projects of computer vision systems for automatic traffic control]. *Avtomatika, svyaz', informatika* [Automation, Communications, Informatics]. 2023, Iss. 3, pp. 21–24. DOI: 10.34649/AT.2023.3.3.003. (In Russian)
7. Ozerov A. V., Marshova A. S. Tekhnicheskoe zrenie v sostave sistem avtomaticheskogo upravleniya dvizheniem poezdov [Computer vision as part of automatic train traffic control systems]. *Perspektivnye informatsionnye tekhnologii (PIT 2022): trudy Mezhdunarodnoy nauchno-tekhnicheskoy konferentsii*, 18–21 *aprelya 2022 goda, pod red. S. A. Prokhorova* [Advanced Information Technologies (PIT 2022): Proc. Int. Sci.-Tech. Conf., Samara, 18–21 April 2022, ed. by S. A. Prokhorov]. Samara: Izdatel'stvo Samarskogo nauchnogo tsentra RAN Publ., 2022, pp. 201–205. (In Russian)
8. Lisanti G., Karaman S., Pezzatini D. A Multi-Camera Image Processing and Visualization System for Train Safety Assessment. arXiv preprint, 2015, arXiv:1507.07815. DOI: 10.48550/arXiv.1507.07815.
9. Saritas M. M. et al. Railway Track Fault Detection with ResNet Deep Learning Models. 2023 International Conference on Intelligent Systems and New Applications (ICISNA'23), 2023.
10. Sun Kh., Chzhuan S., Kostrov A. A. Obzor sovremennykh sistem tekhnicheskogo zreniya, primenyaemykh v transportnoy otrasli [Review of modern computer vision systems used in the transport industry]. *Sovremennye naukoemkie tekhnologii* [Modern High-Tech Technologies]. 2024, Iss. 9, pp. 69–73. DOI: 10.17513/snt.40150. (In Russian)
11. Goodfellow I., Shlens J., Szegedy C. Explaining and Harnessing Adversarial Examples. arXiv preprint, 2015 (submitted 20 Dec 2014, revised 20 Mar 2015), arXiv:1412.6572. DOI: 10.48550/arXiv.1412.6572.
12. Grigorenko A. G., Vasil'ev N. A., Sitdikov D. S. Obzor metodov zashchity ot adversial'noy ataki One Pixel v sistemakh mashinnogo obucheniya [Review of methods of protection against the One Pixel adversarial attack in machine learning systems]. *Sistemy intellektual'nogo upravleniya i iskusstvennyy intellekt: teoriya i praktika: sbornik trudov II natsional'noy nauchno-prakticheskoy konferentsii, Sankt-Peterburg, 27 iyunya 2024 goda* [Intelligent control systems and artificial intelligence: theory and practice: collected papers of the II national scientific and practical conference, St. Petersburg, June 27, 2024]. St. Petersburg: Federal'noe gosudarstvennoe byudzhetnoe obrazovatel'noe uchrezhdenie vysshego obrazovaniya Gosudarstvennyy universitet morskogo i rechnogo flota im. admirala S. O. Makarova Publ., 2024, pp. 30–36. (In Russian)
13. Kh'yuber P. *Robastnaya statistika* [Robust statistics]. Moscow: Mir Publ.,1984. (In Russian)
14. Goodfellow I. J., Shlens J., Szegedy C. Making machine learning robust against adversarial inputs. Communications of the ACM, 2018, vol. 61, Iss. 7, pp. 56–66. DOI: 10.1145/3134599.
15. Vasilenko M. S., Kopyrin A. S. *Algoritm mashinnogo obucheniya dlya detektirovaniya vybrosov i anomaliy* [A machine learning algorithm for detecting outliers and anomalies]. Modeling of Artificial Intelligence, 2019, Iss. 6-1, pp. 13–18. DOI: 10.13187/mai.2019.1.13. (In Russian)
16. Legashev L. V., Grishina L. S. Metodika postroeniya ustoychivoy sistemy zashchity na osnove sostyazatel'nogo mashinnogo obucheniya v besprovodnykh setyakh 6G [Methodology for constructing a robust security system based on adversarial machine learning in 6G wireless networks]. *Voprosy kiberbezopasnosti* [Cybersecurity Issues]. 2023, Iss. 2(54), pp. 99–108. DOI: 10.21681/2311-3456-2023-2-99-108. (In Russian)
17. Hou X. et al. High-Speed Rail Operating Environment Recognition Based on Neural Network and Adversarial Training. 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), IEEE, 2019, pp. 840–847. DOI: 10.1109/ICTAI.2019.00120.
18. Goldobin I. A., Klimova E. I. Vliyanie shumov na algoritmy tsifrovoy obrabotki izobrazheniy [The Impact of Noise on Digital Image Processing Algorithms]. *Aktual'nye voprosy razvitiya sovremennoy tsifrovoy sredy: sbornik statey po materialam nauchno-tekhnicheskoy konferentsii molodykh uchenykh, Moskva, 14–16 aprelya 2021 goda* [Current Issues in the Development of the Modern Digital Environment: A Collection of Articles Based on the Proceedings of the Scientific and Technical Conference of Young Scientists, Moscow, April 14–16, 2021]. Volgograd: Sirius Publ., 2021, pp. 396–402. (In Russian)
19. Kotenko I. V., Saenko I. B., Lauta O. S. et al. Ataki i metody zashchity v sistemakh mashinnogo obucheniya: analiz sovremennykh issledovaniy [Attacks and Defense Methods in Machine Learning Systems: An Analysis of Modern Research]. *Voprosy kiberbezopasnosti* [Cybersecurity Issues]. 2024, Iss. 1(59), pp. 24–37. DOI: 10.21681/2311-2024-1-24-37. (In Russian)
20. Kostyumov V. V. *Obzor i sistematizatsiya atak ukloneniem na modeli komp'yuternogo zreniya* [Review and systematization of evasion attacks on computer vision models]. International Journal of Open Information Technologies, 2022, vol. 10, Iss. 10, pp. 11–20. (In Russian)
21. Potapov A. K., Sidorenko V. G. Vulnerabilities of Artificial Intelligence Systems. 2024 International Conference "Quality Management, Transport and Information Security, Information Technologies" (QM&TIS&IT), IEEE, 2024, pp. 84–87. DOI: 10.1109/QMTISIT63393.2024.10762915.
22. Grachev Ya. L., Sidorenko V. G. Ispol'zovanie kachestvennykh kharakteristik izobrazheniya dlya kompleksnogo stegoanaliza [Using qualitative image characteristics for complex steganalysis]. *Nadezhnost'* [Reliability]. 2025, vol. 25, Iss. 1, pp. 67–74. DOI: 10.21683/1729-2646-2025-25-1-67-74. (In Russian)
23. Orlov S. P., Efimushkin N. A., Efimushkina N. V. Glubokaya neyronnaya set' dlya diagnostiki elementov zheleznodorozhnogo rel'sovogo puti [Deep neural network for diagnostics of railway track elements]. *Vestnik Samarskogo gosudarstvennogo tekhnicheskogo universiteta. Seriya: Tekhnicheskie nauki* [Bulletin of Samara State Technical University. Series: Technical Sciences]. 2022, vol. 30, Iss. 1(73), pp. 63–74. DOI: 10.14498/tech.2022.1.4. (In Russian)
24. Fedorov V. A. Obnaruzhenie zheleznodorozhnoy infrastruktury na osnove YOLOv8 s ispol'zovaniem neyronnogo protsessora [YOLOv8-Based Railway Infrastructure Detection Using a Neural Processor]. *Doklady Rossiyskoy akademii nauk. Matematika, informatika, protsessy upravleniya* [Doklady RAS. Mathematics, Informatics, Control Processes]. 2024, vol. 520, Iss. S2, pp. 49–56. DOI: 10.31857/S2686954324700371. (In Russian)
25. Eunus S. I., Hossain S., A. Ridwan E. M., Adnan A. et al. ECARRNet: An Efficient LSTM-Based Ensembled Deep Neural Network Architecture for Railway Fault Detection. AI, 2024, vol. 5, Iss. 2, pp. 482–503. DOI: 10.3390/ai5020024.
26. Goodfellow I., Bengio Y., Courville A. Deep Learning. Cambridge, MA: MIT Press, 2016. DOI: 10.1007/s10710-017-9314-z.
27. Han K., WangY., Chen H. A Survey on Visual Transformer. arXiv preprint, 2020. DOI: 10.48550/arXiv.2012.12556.
28. Powers D. M. W. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. Journal of Machine Learning Technologies, 2011, vol. 2, Iss. 1, pp. 37–63. DOI: 10.48550/arXiv.2010.16061.