

УДК 004.8+004.048

МОДЕЛИ И МЕТОДЫ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ДЕТЕКТИРОВАНИЯ ОБЪЕКТОВ В ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМАХ ОБРАБОТКИ ВИЗУАЛЬНЫХ ДАННЫХ



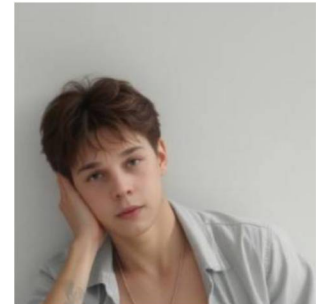
А.А. Навроцкий
Доцент кафедры
информационных
технологий
автоматизированных
систем БГУИР,
кандидат физико-
математических
наук, доцент
dekfitu@bsuir.by



Л.Ю. Шилин
Профессор кафедры
информационных
технологий
автоматизированных
систем БГУИР,
кандидат
технических наук,
доктор технических
наук
shilin@bsuir.by



Д.Д. Хмыз
Ассистент кафедры
вычислительных
методов и
программирования
БГУИР, магистрант
кафедры ИТАС
d.khmyz@bsuir.by



К.Д. Сморщок
Инженер-
программист,
ООО «Эфкон
диджитал»
s.kirill150282@mail.ru

А.А. Навроцкий

Окончил Минский радиотехнический институт. Область научных интересов связана с разработкой систем технического зрения и технологиями программирования.

Л.Ю. Шилин

Окончил Минский радиотехнический институт, специальность «Автоматика и телемеханика». Область научных интересов связана с анализом и синтезом цифровых систем автоматизированного управления.

Д.Д. Хмыз

Окончила Белорусский национальный технический университет. Область научных интересов связана с технологиями программирования и разработкой систем технического зрения.

К.Д. Сморщок

Область научных интересов связана с технологиями программирования.

Аннотация. В работе представлено исследование и практическая реализация нейросетевой системы детектирования объектов транспортной инфраструктуры на основе современной архитектуры Ultralytics YOLO26. Рассмотрены методы глубокого обучения для обработки визуальных данных в задачах реального времени, особенности оптимизации параметров сети с использованием модифицированного алгоритма MuSGD, а также комплексная оценка качества модели по метрикам IoU, Precision, Recall, F1 и mAP. Полученные результаты ($mAP@0.5 = 0.768$, Precision = 0.93, Recall = 0.84) подтверждают эффективность предложенного подхода для применения в интеллектуальных информационных системах мониторинга.

Ключевые слова: глубокое обучение, детектирование объектов, обработка визуальных данных, YOLO, интеллектуальные информационные системы, оптимизация.

Введение. Анализ визуальной информации с помощью методов глубокого обучения становится ключевым направлением в развитии интеллектуальных систем. Высокая скорость и точность распознавания объектов в реальном времени открывают широкие возможности для автоматизации мониторинга, управления и поддержки принятия решений.

При этом важным фактором является оптимизация нейросетевых архитектур с учётом ограничений вычислительных ресурсов и требований к минимальной задержке обработки.

Основная часть. Современные алгоритмы детектирования объектов активно развиваются, выходя за рамки традиционных схем и позволяя достигать новых уровней производительности и стабильности. Современные информационные технологии в сфере обработки данных активно используют методы глубокого обучения для анализа визуальной информации.

В задачах интеллектуального мониторинга, автоматизированного управления и систем поддержки принятия решений особое значение приобретает корректное и оперативное детектирование объектов на изображениях и видеопотоках. Такие системы функционируют в условиях ограниченных вычислительных ресурсов и требований к реальному времени обработки, что определяет необходимость использования эффективных архитектур нейронных сетей.

Эволюция методов object detection прошла путь от двухступенчатых детекторов (Faster R-CNN) [1] к одношаговым архитектурам семейства YOLO [2], обеспечивающим баланс между точностью и скоростью обработки.

Современные end-to-end модели, включая трансформерные подходы (DETR) [3], минимизируют необходимость постобработки и этапа подавления немаксимумов (NMS), что снижает вычислительную нагрузку и повышает стабильность инференса.

В настоящей работе применена архитектура Ultralytics YOLO26 [4], относящаяся к одношаговым детекторам нового поколения с головой типа one-to-one.

Исключение NMS позволяет обеспечить предсказуемое время обработки кадра и уменьшить латентность, что критично для информационных систем реального времени.

Обучение модели выполнялось с использованием стратегии трансферного обучения на основе предобученных весов датасета COCO [5].

Такой подход позволяет использовать универсальные визуальные признаки низкого и среднего уровня (контуры, текстуры, геометрические структуры), адаптируя их к специализированной предметной области.

Оптимизация параметров сети осуществлялась с применением гибридного алгоритма MuSGD. Классический стохастический градиентный спуск реализуется по формуле:

$$wt+1=wt-\eta L(wt), \quad (1)$$

где η – скорость обучения,

$L(wt)$ – градиент функции потерь.

Модификация MuSGD включает нормализацию градиента с учётом оценки его дисперсии:

$$vt=\mu vt-1+L(wt)st+\varepsilon. \quad (2)$$

Такой подход стабилизирует процесс обучения, уменьшает осцилляции и ускоряет достижение минимума функции потерь, что особенно важно при высокой размерности параметрического пространства.

Для оценки качества модели использовался комплекс метрик, традиционно применяемых в задачах детектирования объектов.

Коэффициент перекрытия IoU (Intersection over Union) определяется как отношение площади пересечения предсказанной и истинной рамок к площади их объединения:

$$IoU=\frac{Area_{intersection}}{Area_{union}}. \quad (3)$$

Точность (Precision) рассчитывается как:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (4)$$

а полнота (Recall):

$$\text{Recall} = \frac{TP}{TP + FN}. \quad (5)$$

Гармоническое среднее этих показателей определяется через F1-метрику:

$$F1 = 2 \frac{P \cdot R}{P + R}. \quad (6)$$

Интегральной характеристикой выступает средняя точность (mAP), вычисляемая как среднее значение Average Precision по всем классам.

В результате обучения получены следующие показатели качества:

$$mAP@0.5 = 0.768, \text{ Precision} = 0.93, \text{ Recall} = 0.84, F1 = 0.62.$$

Высокое значение Precision свидетельствует о низкой вероятности ложных срабатываний, а Recall = 0.84 подтверждает способность модели обнаруживать большинство целевых объектов. Значение mAP@0.5 характеризует устойчивость модели при варьировании порога уверенности.

Кривая Precision–Recall для модели YOLO26 представлена на рисунке 1.

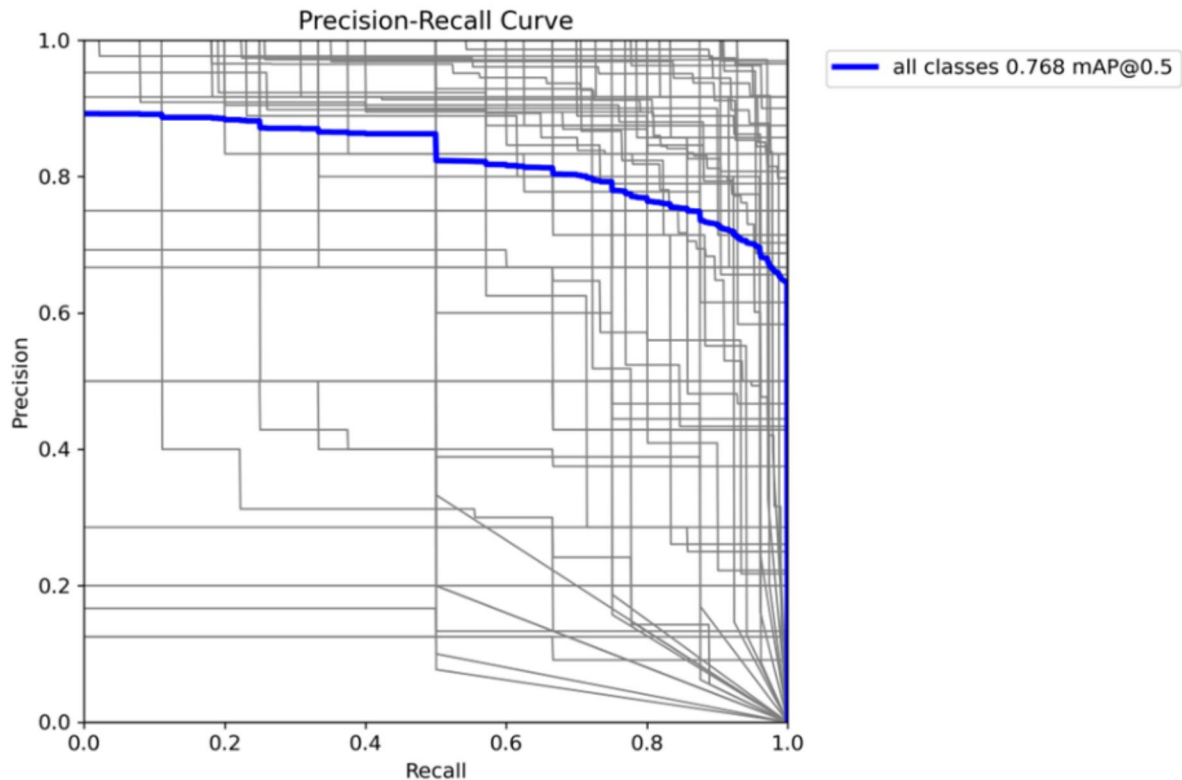


Рисунок 1 Кривая Precision–Recall для модели YOLO26

Особое внимание уделялось анализу дисбаланса классов (long-tail распределения), характерного для реальных датасетов [6]. Наличие редких категорий объектов оказывает влияние на снижение F1-метрики, однако интегральные показатели mAP демонстрируют достаточную обобщающую способность модели.

Заключение. Проведённое исследование демонстрирует эффективность современных методов глубокого обучения для задач автоматизированного детектирования объектов в интеллектуальных информационных системах.

Использование архитектуры YOLO26 и оптимизатора MuSGD позволяет обеспечить баланс между точностью и вычислительной эффективностью, что делает предложенный подход перспективным для внедрения в системы реального времени.

Список литературы

- [1] Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // IEEE TPAMI. – 2017.
- [2] Redmon J., Farhadi A. YOLOv3: An Incremental Improvement // arXiv:1804.02767. – 2018.
- [3] Carion N. et al. End-to-End Object Detection with Transformers // ECCV. – 2020.
- [4] Ultralytics. Ultralytics YOLO Documentation. – 2024.
- [5] Lin T.-Y. et al. Microsoft COCO: Common Objects in Context // ECCV. – 2014.
- [6] Zhang Y. et al. Deep Long-Tail Learning: A Survey // IEEE TPAMI. – 2023.

Авторский вклад

Анатолий Александрович Навроцкий – постановка задачи исследования, формулирование основных алгоритмов, разработка структуры статьи.

Леонид Юрьевич Шилин – постановка задачи исследования и формулировка гипотез.

Дарья Дмитриевна Хмыз – разработка алгоритмов работы системы, сбор и систематизация фактических данных, написание статьи.

Кирилл Дмитриевич Сморщок – обучение интеллектуальной системы обработки.

DEEP LEARNING MODELS AND METHODS FOR OBJECT DETECTION IN INTELLIGENT VISUAL DATA PROCESSING SYSTEMS

A.A. Navrotsky

*Associate Professor,
Department of Information
Technologies for Automated
Systems, BSUIR, PhD in
Physics and Mathematics,
Associate Professor*

L. Yu. Shilin

*Professor, Department of
Information Technologies
and Automated Systems,
BSUI), Candidate of
Technical Sciences,
Doctor of Technical
Sciences*

D.D. Khmyz

*Assistant Professor,
Department of
Computational Methods
and Programming, BSUIR,
Master's Student,
Department of ITAS*

K.D. Smorshchok

*Student, group
320602, BSUIR*

Abstract. The paper presents a study and practical implementation of a neural network-based object detection system for transport infrastructure using the modern Ultralytics YOLO26 architecture. Methods of deep learning for visual data processing in real-time applications are considered, along with network parameter optimization using the modified MuSGD algorithm. A comprehensive evaluation of model performance is conducted using IoU, Precision, Recall, F1-score, and mAP metrics. The obtained results ($mAP@0.5 = 0.768$, Precision = 0.93, Recall = 0.84) confirm the effectiveness of the proposed approach for deployment in intelligent information monitoring systems.

Keywords: deep learning, object detection, visual data processing, YOLO, intelligent information systems, optimization.