

## МОНИТОРИНГ АКТИВНОСТИ ЛЮДЕЙ В МЕСТАХ МАССОВОГО СКОПЛЕНИЯ ПО ДАННЫМ ВИДЕОНАБЛЮДЕНИЯ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

Шалыгин В. А.<sup>1</sup>

Научный руководитель – к.т.н., доцент Ярышев С. Н.<sup>1</sup>

<sup>1</sup> Университет ИТМО

vashalygin@itmo.ru

ysn63@mail.ru

Работа выполнена в рамках темы НИР №625111 «Алгоритмическая прозрачность и предиктивные модели развития AI-технологий: от базы знаний к навигатору решений».

### Введение

С развитием городской инфраструктуры и ростом плотности населения задачи автоматизированного анализа видеопотоков приобретают критическую важность для обеспечения общественной безопасности и мониторинга людских потоков [1, 2]. Системы мониторинга, способные в реальном времени детектировать и отслеживать людей в плотных скоплениях, сталкиваются с рядом проблем: частые окклюзии, высокая схожесть визуальных признаков (например, однотипная одежда), динамические изменения поз и необходимость высокоскоростной обработки [2, 3]. В связи с этим цель настоящей работы – разработка алгоритма позиционирования и трекинга, устойчивых к характерным для плотных сцен искажениям и помехам, – становится важной научно-технической задачей.

### Основная часть

Отслеживание множества объектов (Multi-Object Tracking, MOT) — это процесс последовательного обнаружения и сопровождения объектов по кадрам видеопоследовательности [1,3]. Современные подходы преимущественно реализуются по принципу «трекинг-через-детекцию», где качество финального трекинга напрямую зависит от стабильности работы детектора. В данной работе в качестве детекторов рассматриваются и сравниваются архитектуры семейства YOLO (YOLOv5, YOLOv8, YOLOv9) [4, 5, 6], а для ассоциации объектов между кадрами используется трекер, основанный на фильтре Калмана [7].

Детекторы были дообучены на собранном в рамках работы датасете, включающем в себя различные типы скопления людей. Основой для его создания послужили 20 видеопотоков, каждый из которых был разбит на 3600 изображений, которые отражают три сценария движения: условно структурированное, неструктурированное и строго структурированное. Это позволяет всесторонне оценить работу алгоритмов в условиях, максимально приближенных к реальным прикладным задачам.

В рамках работы использован пайплайн, объединяющий детекторы YOLO различных версий с трекером DeepSORT. Первоначальный анализ работы детекторов показал, что архитектуры YOLOv5 и YOLOv8 демонстрируют сбалансированные показатели точности (precision) и полноты (recall) на сценах со средней плотностью. Однако для создания устойчивого трекера в условиях экстремально плотных скоплений потребовался более глубокий анализ влияния архитектурных особенностей детекторов на стабильность трекинга. Для оценки качества трекинга применялись следующие метрики: точность отслеживания множества объектов (MOTA), прецизионность локализации (MOTP) и количество переключений идентификаторов (IDsw).

Результаты промежуточных экспериментов продемонстрировали, что YOLOv5-

small в связке с DeepSORT обеспечивает наилучший баланс между точностью и стабильностью трекинга. Максимальные показатели MOTA составили 0.4 при 60 переключениях идентификаторов, что свидетельствует о высокой согласованности детекций. YOLOv8 показал результаты MOTA 0.29, IDsw 49, однако продемонстрировал лучшую детекцию мелких объектов. YOLOv9, несмотря на более современные архитектурные решения, показал нестабильные результаты: на сценах, содержащих экстремально плотное неструктурированное движение, значение MOTA оказалось отрицательным, что связано с превышением количества ошибок над числом объектов в эталонной разметке.

### **Выводы**

Разработан и экспериментально протестирован алгоритм позиционирования и трекинга людей на основе комбинации детекторов семейства YOLO и трекера DeepSORT, обученный на собранном корпусе данных. Установлено, что ключевым фактором успешного трекинга в плотных сценах является не максимальная точность детекции, а стабильность и согласованность выдаваемых ограничивающих рамок. YOLOv5 показала себя как наиболее сбалансированное решение для интеграции с DeepSORT, обеспечивая надежное отслеживание с минимальным количеством переключений идентификаторов. Результаты тестирования на эталонных наборах данных MOT17 и MOT20 также подтверждают эффективность предложенного подхода и позволяют определить направления дальнейшего улучшения алгоритмов для работы в условиях высоких скоплений.

### **Литература**

1. Gajjar V., Gurnani A., Khandhediya Y. Human detection and tracking for video surveillance: A cognitive science approach //Proceedings of the IEEE international conference on computer vision workshops. – 2017. – С. 2805-2809.
2. Ravindran R., Santora M. J., Jamali M. M. Multi-object detection and tracking, based on DNN, for autonomous vehicles: A review //IEEE Sensors Journal. – 2020. – Т. 21. – №. 5. – С. 5668-5677.
3. Brunetti A. et al. Computer vision and deep learning techniques for pedestrian detection and tracking: A survey //Neurocomputing. – 2018. – Т. 300. – С. 17-33.
4. Liu H. et al. Sf-yolov5: A lightweight small object detection algorithm based on improved feature fusion mode //Sensors. – 2022. – Т. 22. – №. 15. – С. 5817.
5. Talaat F. M., ZainEldin H. An improved fire detection approach based on YOLO-v8 for smart cities //Neural computing and applications. – 2023. – Т. 35. – №. 28. – С. 20939-20954.
6. Wang, C.Y.; Yeh, I.H.; Liao, H.Y.M. Yolov9: Learning what you want to learn using programmable gradient information. arXiv 2024, arXiv:2402.13616.
7. Du Y. et al. Strongsort: Make deepsort great again //IEEE Transactions on Multimedia. – 2023. – Т. 25. – С. 8725-8737.