

УДК 004.89

ОСОБЕННОСТИ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ ОБНАРУЖЕНИЯ ЧЕЛОВЕКА В БЕЗОПАСНЫХ ЗОНАХ

Наурузов У.М. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Ключев А.О.
(ИТМО)

Введение. В настоящее время, в век активного развития искусственного интеллекта и, в частности, моделей нейронных сетей, связанных с компьютерным зрением и применением их в задачи обнаружения различного рода объектов и людей в охраняемой зоне, возникает необходимость поиска эффективного решения этой задачи. Такого вида программное обеспечение должно воплощать в себе высокую производительность и точность, а также возможность работы, что называется “runtime”, то есть работать в режиме реального времени. Кроме этого, такие программы находят обширное применение, начиная с городской среды и заканчивая режимными объектами. При этом, такое программное обеспечение должно включать в себя применение различных методик и алгоритмов для оптимизации всех этапов работы программы. В целом, было произведено исследование имеющихся методов обнаружения человека в безопасных зонах.

Основная часть. Для обнаружения человека используются различные модели сверхточных нейронных сетей, таких как Faster R-CNN, SSD, Mask R-CNN и семейство моделей YOLOv[1]. Особое внимание стоит уделить моделям семейства YOLOv. За период существования данного вида моделей было выпущено несколько версий, последней из которых является версия YOLOv8. Модели семейства YOLOv зарекомендовали себя как отличные системы обнаружения объектов в режиме реального времени для робототехники, беспилотных автомобилей и приложений видеомониторинга. В моделях YOLOv используется “Non-Maximum Suppression” (далее NMS) - это метод постобработки, используемый в алгоритмах обнаружения объектов для уменьшения количества перекрывающихся ограничивающих рамок и улучшения общего качества обнаружения. Алгоритмы обнаружения объектов обычно генерируют несколько ограничивающих рамок вокруг одного и того же объекта с разными показателями достоверности. NMS отфильтровывает избыточные и нерелевантные ограничивающие рамки, сохраняя только наиболее точные[2]. При этом модели семейства YOLOv в частности YOLOv8 изначально обучена на датасете Microsoft COCO и осуществляет детекцию 80 категорий объектов и, кроме этого, довольно хорошо и успешно поддаются дообучению, что позволяет обучить модель обнаруживать необходимые для программного обеспечения объекты. Подготовка датасета для дообучения модели - это большой процесс, при котором идет подбор подходящих фотографий и ручное выделение координат рамки вокруг объектов, которые необходимо детектировать в процессе работы модели в программном обеспечении, а также разумное применение методов аугментации для получения большего количества фотографии. Также при формировании датасета стоит учитывать высоту и угол наклона камеры, с которым программное обеспечение будет получать видеопоток для получения приемлемой точности детекции и корректной работы программы.

Следующий этап - это выбор алгоритма построения границ безопасной зоны. Первым вопросом, который стоит решить, является в каком формате использовать модель нейронных сетей, есть форматы “pt”, “onnx”, “engine” и другие, и для каждого формата имеется своя специфика работы и алгоритм обработки. Наиболее предпочтительным является формат TensorRT - ‘engine’, который использует механизм, при котором наиболее слабые нейроны “обрезаются”, тем самым модель оптимизируется, что позволяет увеличить скорость работы модели.

Следующим вопросом, который необходимо решить, выбор модели нейронных сетей. Здесь предпочтение стоит отдать наиболее новейшей модели семейства YOLOv, то есть

YOLOv8, которая вобрала в себя наиболее лучшие практики и механизмы предыдущих версий.

Заключительным вопросом, является выбор алгоритма определения и построения границ безопасной зоны. Первым из вариантов является обучение модели определению объектов, обычно использующихся для определения границ различных зон, таких как конусы, после чего соединения точек середины нижней части “bounding box”, то есть выделенной границы объекта в данном случае и построения на их основе границ безопасной зоны. Второй вариант предусматривает определение точек вручную на полученном с камеры видеопотоке и построении на основе предоставленных точек безопасной зоны. Построение границ безопасной зоны начинается, когда определены не менее трех точек.

Выводы. Проведен анализ особенностей разработки программного обеспечения для обнаружения человека в безопасных зонах и разработана методика разработки такого вида программ.

Список использованных источников:

1. Fang W., Peter E.D.L., Luo H., Ding L. Computer vision for behaviour-based safety in construction: A review and future directions. - 2019. URL: <https://www.sciencedirect.com/science/article/abs/pii/S1474034619305531>, дата последнего посещения 10.01.2024.
2. Terven J., Cordova-Esparza D.-M., Romero-Gonzalez J.-A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. URL: https://www.researchgate.net/publication/375783101_A_Comprehensive_Review_of_YOLO_Architectures_in_Computer_Vision_From_YOLOv1_to_YOLOv8_and_YOLO-NAS, дата последнего посещения 15.01.2024.