

РЕШЕНИЕ ЗАДАЧИ РАСПОЗНАВАНИЯ НОГ ПЕШЕХОДОВ С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ YOLOV4

*Т. Е. Мамонова, к.т.н., доцент ОАР ИШИТР,
Я. К. Завгородний, студент гр. 8ЕМ01
Томский политехнический университет
E-mail: ykz2@tpu.ru*

Введение

В последнее время обнаружение объектов на основе сверточных нейронных сетей стало популярной темой исследований в области компьютерного зрения. Выделение признаков на исходных изображениях могут быть проведены с помощью многослойной операции свертки, а положение объекта на изображении предсказываются с помощью специальных алгоритмов. Результаты этих исследований активно применяют в мехатронике и робототехнике в задачах навигации мобильных роботов, управления промышленными манипуляторами и т.п. [1].

Целью данной работы является анализ работы сверточной нейронной сети YOLOv4 для распознавания ног пешеходов с дальнейшим применением результатов для решения задачи навигации мобильных роботов (например, роботов-уборщиков в аэропортах, торговых центрах).

Применение сверточных нейронных сетей в задаче навигации мобильного робота

Обнаружение и отслеживание человека необходимы для сервисных автономных мобильных роботов (АМР), поскольку робот часто делит свое рабочее пространство и тесно взаимодействует с людьми. Например, разработка FLOBOT использует 3D-лидар, RGB-камеру и 2D-лидар для обнаружения и отслеживания человека [2].

Данные, полученные с камеры, позволяют организовать правильную работу алгоритмов навигации мобильных роботов. Для этого необходимо добиться высокой точности обучения сверточной нейронной сети. Разработанная система с использованием алгоритма распознавания ног пешеходов дает возможность определить траекторию движения человека. На рисунке 1 представлена функциональная схема работы АМР с выделенной схемой исследовательской части.

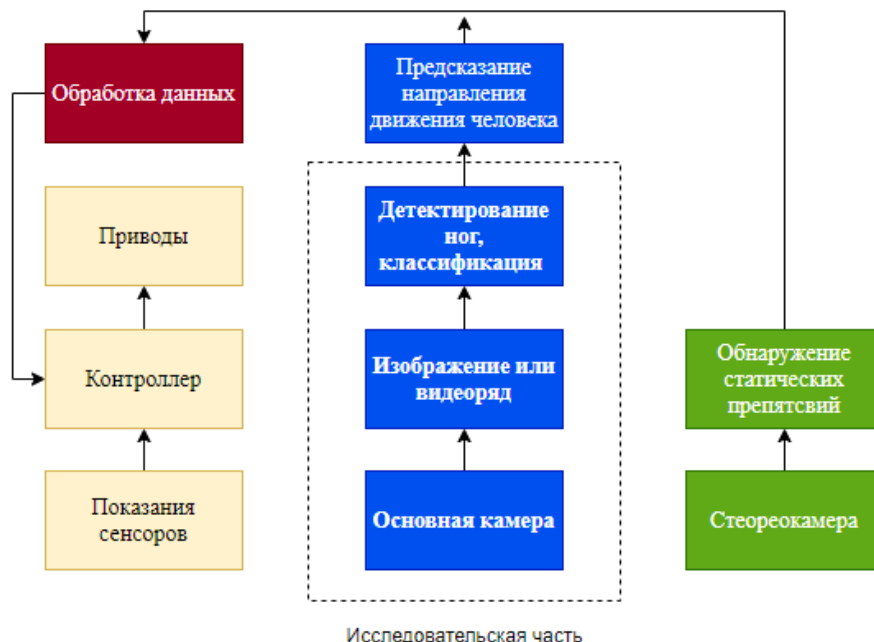


Рис. 1. Функциональная схема работы АМР

На сегодняшний день нейронная сеть архитектуры YOLO, является одной из самых эффективных нейросетей для детектирования объектов [3–4].

Описание нейронной сети YOLO

Многие разработчики, в том числе и проектирующие системы детектирования статических и динамических препятствий для AMP, используют модификации YOLO для своих задач [4–6].

Процесс детектирования объектов YOLO-сетью состоит из следующих шагов. Сначала форматируется исходное изображение (обычно до формата 416×416 пикселей). Далее на изображение накладывается сетка заранее выбранного размера, затем изображение подается на вход сети. В ходе работы сети выделяются «регионы-кандидаты», где высока доля вероятности нахождения искомого объекта. Далее изображение сжимается на слое pooling и по нему определяется класс объекта. Точность каждого прямоугольника детектирования определяется с помощью метрики IoU [6].

Подготовка данных

Важным этапом обучения YOLO-сети является подготовка аннотаций – разметке данных с целью указания интересующих нас объектов. Разметка данных проводилась с помощью программы labelImg-master, имеющей открытый исходный код [7]. Для обучения было подготовлено 326 изображений с 2 264 аннотациями по четырем классам с долей в 25 %.

Обучение нейронной сети YOLO-4

Для обучения нейронных сетей в данной работе использовались мощности ГП Nvidia Tesla T4, которые предоставляет сервис Google Collab [8]. Результат обучения показан на рисунке 2. Лучший из полученных результатов исследования влияния различных параметров на результат обучения нейронной сети представлен в таблице 1.

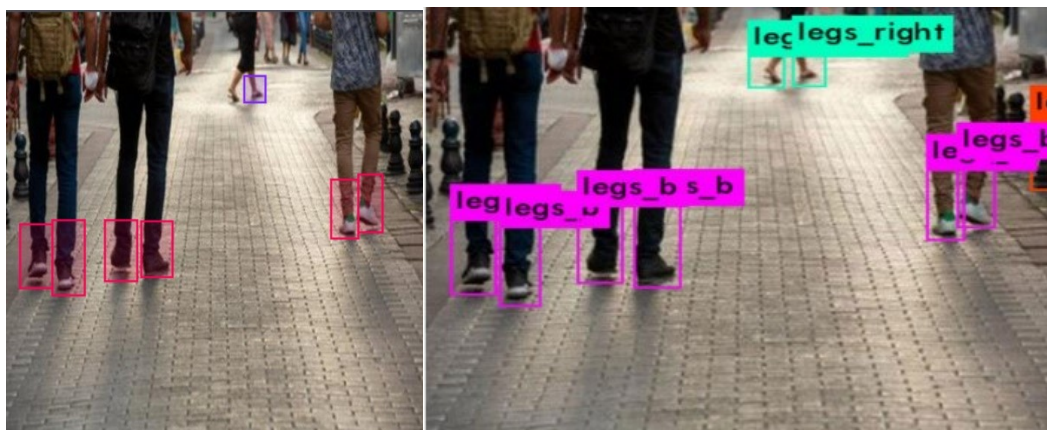


Рис. 2. Сравнение результатов на валидационной выборке с изначальными аннотациями

Таблица 1. Результат обучения

Точность, %	Количество итераций (время обучения 6 часов 14 минут)								
	1000	1100	1200	1300	1400	1500	1700	1900	2000
Общая	62,86	68,63	72,4	75,95	73,94	76,89	79,6	80,15	79,72
Класс 1	71,43	72,26	72,71	73,52	73,96	74,97	73,29	74,44	73,11
Класс 2	63,30	65,46	67,8	74,25	67,51	73,81	72,8	75,46	75,24
Класс 3	63,50	74,67	76,54	81,82	78,06	81,52	86,58	87,71	87,51
Класс 4	53,22	62,14	72,56	74,20	76,23	77,25	85,71	82,99	83,01
IoU	37,32	43,61	45,22	46	48,33	52,78	50,58	53,73	55,35

Как видим, создание набора данных является важным параметром и сбалансированность классов приводит к высокой точности обучения. В этом случае удалось получить лучшую точность в 80,15 % после 1900-й итерации. С увеличением числа итераций результат обучения сети YOLO улучшается.

Заключение

Данная работа продемонстрировала возможности нейронной сети архитектуры YOLOv4 в задаче распознавания ног пешеходов. Полученные результаты можно применять для предсказания

направления движения пешеходов для решения задач навигации автономного мобильного робота. Полученная точность приемлема для данного количества итераций (более 1900). Для улучшения точности необходимо расширять набор данных или увеличить количество итераций, что негативно скажется на скорости обучения.

Список использованных источников

1. Model-based Pedestrian Trajectory Prediction using Environmental Sensor for Mobile Robots Navigation. Haruka Tonoki, Ayanori Yorozu, Masaki Takahashi. International Journal of Advanced Computer Science and Applications, Vol. 8, No. 2, 2017, pp. 2–12.
2. Ögren P, Leonard NE (2005) A convergent dynamic window approach to obstacle avoidance. IEEE Trans Robot 21(2): pp. 188–195.
3. Human Leg Detection from Depth Sensing, Andrei Lucian, Andreea Sandu, Radu Orghidan. Conference: 2018 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR), pp. 2–6.
4. T. Dean, M. Ruzon, M. Segal, J. Shlens, S. Vijayanarasimhan, J. Yagnik, et al. Fast, accurate detection of 100,000 object classes on a single machine. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pp. 18–23.
5. D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, с. 2155–2162. IEEE, 2014
6. You Only Look Once: Unified, Real-Time Object Detection. Joseph Redmon, Santosh Divvalay, Ross Girshick, Ali Farhadi, Университет Вашингтона. – 2016. – pp. 3–10.
7. Программа labellmg. Пакет программ на Github. [Электронный ресурс]. – URL: <https://github.com/tzutalin/labellmg> (дата обращения: 16.02.2022).
8. Google collab. [Электронный ресурс]. – URL: <https://colab.research.google.com> (дата обращения: 18.02.2022).