

ПРИМЕНЕНИЕ КАСКАДА НЕЙРОННЫХ СЕТЕЙ ДЛЯ БЫСТРОГО ТРЕКИНГА ОБЪЕКТОВ В ВИДЕОПОТОКЕ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Базаревский В.Э.

Бранцевич П.Ю. – канд. техн. наук, доцент

В настоящее время задачи, связанные с трекингом объектов, становятся актуальнее с каждым днем. В частности, подобные задачи возникают при реализации идеи “умного” магазина, в котором система должна анализировать поведение клиентов в помещении, строить карты перемещения клиентов и при выходе из магазина автоматически формировать счет. При решении подобной задачи можно столкнуться с проблемой классификации объектов на изображении, с проблемой определения местоположения объекта на изображении и с проблемой идентификации клиентов. Задача такого типа может быть разделена на составные части и решена комплексом методов.

Трекинг объектов должен быть быстрым, точным и способным распознавать самые разнообразные объекты. Для таких целей часто используются R-CNN (Region-based Convolutional Networks). R-CNN (Region-based Convolutional Neural Networks) – это сверточная нейронная сеть, которая позволяет определять местоположение и размер объектов на изображении. Такая сеть в процессе обработки делит изображение на регионы и за счет этого достаточно одного прохода для трекинга нескольких объектов. Есть несколько описанных архитектур, позволяющих это делать:

1. Faster R-CNN;
2. Mask R-CNN;
3. YOLO9000.

Размер потенциального объекта может варьироваться и сразу производится оценка различных как по размеру, так и по форме регионов изображения. В зависимости от поставленной задачи архитектура сети может изменяться. Архитектурой с самым обширным множеством распознаваемых классов является YOLO9000. Под классом в данном случае подразумевается набор общих признаков, присущих определенному типу объектов. Сеть такого типа может классифицировать более 9000 классов [1], что достаточно много для оценки происходящего на изображении, но не для реализации идеи “умного” магазина, т.к. ассортимент каждого магазина, где будет установлена такая система может обновляться и т.к. необходимо получать информацию о перемещениях определенного клиента по магазину. Задача изменения ассортимента решается путем внедрения в упаковку товаров RFID (Radio Frequency Identification) меток. Задача же идентификации определенного клиента должна решаться другим способом. Одной информации, что на изображении находится человек не достаточно. Необходимо использовать метод, позволяющий добавлять клиентам идентификатор, по которому этот же клиент может быть идентифицирован в будущем, при условии появления его в поле зрения другой камеры.

Для решения задачи идентификации клиентов также могут быть использованы нейронные сети. Одним из ярких представителей такого типа сетей является FaceNet. Основной принцип работы этой сети заключается в преобразовании изображения в компактное Евклидово пространство. Расстояние в этом пространстве соответствует подобию лиц. Для обучения такого типа сетей используется оценка совпадающих/не совпадающих изменений цвета частей лица человека. Основными особенностями FaceNet, является то, что для хранения одного лица необходимо всего 128 бит информации и при этом точность идентификации выше 95% [3]. Входными же данными для FaceNet должно быть изображение, на котором находится лишь лицо идентифицируемого человека.

Таким образом, для реализации системы “умного” магазина, необходимо использовать каскад нейронных сетей. В магазине может быть установлено несколько видеокамер высокого разрешения, видеопотоки которых будут вначале обрабатываться нейронной сетью, в основе которой будет лежать YOLO9000. Результатом обработки будут являться координаты и размеры различных объектов. Объекты типа “человек” будут поступать на вход другой нейронной сети, занимающейся непосредственно идентификацией конкретного объекта. Таким образом, после обработки кадра, система будет знать, какой человек где находится. На основании этого можно строить тепловые карты, собирать аналитику о заинтересованности клиента в каком-либо типе товара. Эта информация может быть использована как для маркетинговых исследований, оптимизации работы, повышении уровня продаж, так и в целях безопасности и сохранности имущества владельца магазина.

Список использованных источников:

1. Joseph Redmont, Ali Farhadi YOLO9000: Better, Faster, Stronger
2. Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks
3. Florian Schroff, Dmitry Kalenichenko, James Philbin FaceNet: A Unified Embedding for Face Recognition and Clustering
2. Kaiming He, Georgia Gkioxari, Piotr Dolla, Ross Girshick Mask R-CNN