

УДК 004.855.5

ОБНАРУЖЕНИЕ ЛИЦ С ПОМОЩЬЮ ДЕТЕКТОРА YOLOv3

Аксёнов А.А.

(Санкт-Петербургский институт информатики и автоматизации Российской академии наук,
Университет ИТМО)

Научный руководитель – д.т.н., доцент Карпов А.А.

(Санкт-Петербургский институт информатики и автоматизации Российской академии наук)

В статье рассмотрено использование детектора YOLOv3 для обнаружения лиц с некоторыми улучшениями, приспособленными для решения проблем в задачи обнаружения лиц, среди которых – изменение слоя обнаружения для нахождения мелкогазмерных изображений лиц. В качестве функции потерь используется Softmax, нужная для максимизации разницы между классами признаков и уменьшения размеров этих признаков на слоях обнаружения с целью повышения скорости работы.

Задача обнаружения лиц является первоначальным и особо важным условием в различных исследованиях, таких как распознавание и идентификация лиц, анализ выражения лиц и т.п. В существующих одно- и двухэтапных методах обнаружения объектов, чаще при высокой скорости обнаружения страдает точность, и наоборот. Ярким примером таких детекторов является детектор на основе метода Виолы-Джонса, в котором высокая скорость обнаружения, но довольно низкая точность. Противоположным такому подходу являются детекторы лиц на основе глубокого обучения, они обладают крайне высокой точностью, но с недостатком в виде медленной скоростью детектирования лиц. В то время как, YOLOv3, популярный алгоритм обнаружения объектов, имеет очевидные преимущества как по скорости, так и в точности обнаружения лиц.

Цель работы заключается в проведении экспериментов по обнаружению лиц на специализированной базе данных русского жестового языка TheRusLan с использованием детектора YOLOv3.

На вход подаются видеоданные в двумерном формате с разрешением 1920x1080 пикселей. В качестве детектора лиц выступает YOLOv3, который является усовершенствованной версией архитектуры YOLO и состоит из 106 сверточных слоев, благодаря чему лучше обнаруживает небольшие объекты по сравнению с YOLOv2. Главной особенностью YOLOv3 является то, что на выходе есть три слоя каждый из которых рассчитан на обнаружение объектов разного размера. Отличительной особенностью архитектуры YOLO по сравнению с другими архитектурами состоит в том, что большинство систем применяют свертку несколько раз к разным регионам изображения, в YOLO свертка применяется один раз ко всему изображению сразу. Сеть делит изображение на своеобразную сетку и предсказывает ограничительные рами и вероятности того, что там есть искомый объект для каждого участка.

Таким образом, с помощью детектора YOLOv3 возможно быстро и точно обнаруживать лица, с наличием разного рода помех, например, окклюзия, которая зачастую встречается в используемой специализированной базе данных. Благодаря своей скорости и точности данный детектор может быть использован в задачах биометрии, компьютерном зрении, машинном обучении, автоматических системах распознавания лиц.