

КЛАССИФИКАЦИЯ ПОЛА И ВОЗРАСТА ЛЮДЕЙ В ВИДЕОПОТОКЕ С ПОМОЩЬЮ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ

Семенюта А.В.

НИ ТПУ, ИШИТР, 8ПМ11, e-mail: avs183@tpu.ru

Введение

Как правило, задачу распознавания пола и возраста людей решают с помощью выделения и анализа лиц [1]. Однако на данных с общими условиями съёмки лицо человека может быть распознано далеко не всегда: лицо может быть повёрнуто и может не попасть на камеру, лицо может быть скрыто, условия освещения могут помешать распознаванию лица, лицо может быть не распознано из-за слишком большого расстояния от камеры до человека.

Целью данной работы являлось написание и обучение алгоритма распознавания пола и возраста, использующего вместо изображений лица признаки силуэта и движения (походки) человека. Такие методы могут быть применены в случаях, когда лицо человека распознать не представляется возможным.

Сбор данных

В качестве данных для решения задачи сегментации использовался набор данных UP-S31 [2]. Это специфический набор данных, предназначенный непосредственно для решения задачи сегментации тела человека на изображении и отделении его от остального фона. Исходный набор содержит 8515 размеченных изображений. В силу небольшого количества исходных данных они подвергались аугментации. Итоговое количество изображений – 8515.

В качестве данных для решения задачи классификации использовались наборы данных PETA – Town Centre (6968 изображений) [3] и MARS (1191003 изображения) [4]. Данные формировались следующим образом: сначала происходила фильтрация некорректно размеченных изображений и кадрирование человека на изображении. Затем были составлены треклеты – наборы положений человека на каждом кадре видеопоследовательности. Помимо оригинальных изображений треклетов применялись и аугментированные изображения. В итоге осталось 159 уникальных объектов на PETA – Town Center и 1028 уникальных объектов на MARS.

Описание алгоритма

Основная часть алгоритма работает с результатом трекинга – базой отслеживаемых на видеопотоке объектов, для которых уже было произведено обновление состояния, соответствующее текущему обрабатываемому кадру.

Логика работы алгоритма для одной итерации обработки одного объекта одного кадра видеопотока следующая:

1. Алгоритм определяет принадлежность объекта к классу «Человек». Только для объектов этого класса и возможно определение пола и возраста.
2. Для каждого такого объекта выбирается последнее ограничивающее окно, в рамках которого выполняется сегментация силуэта человека. Данная операция осуществляется с помощью модели сегментации MobileSeg из библиотеки PaddleSeg.
3. Среди набора извлечённых признаков ограничивающих окон объектов выбирается поднабор, описывающий ограниченный по длительности треклет.
4. На основе признаков ограничивающих окон строится GEI изображение – усреднённое изображение бинаризованных силуэтов.
5. Осуществляется вызов модели, предварительно обученной методом переноса обучения, на базе архитектуры ResNet50, для определения пола и возраста человека по его GEI-изображению.
6. На основе всей истории результатов распознавания пола и возраста текущего объекта осуществляется агрегация как выбор чаще всего встречающегося класса, и принимается окончательное решение о значении пола и возраста для объекта.

В общем виде конвейер преобразований данных для одного треклета объекта выглядит так, как показано на рисунке 1.

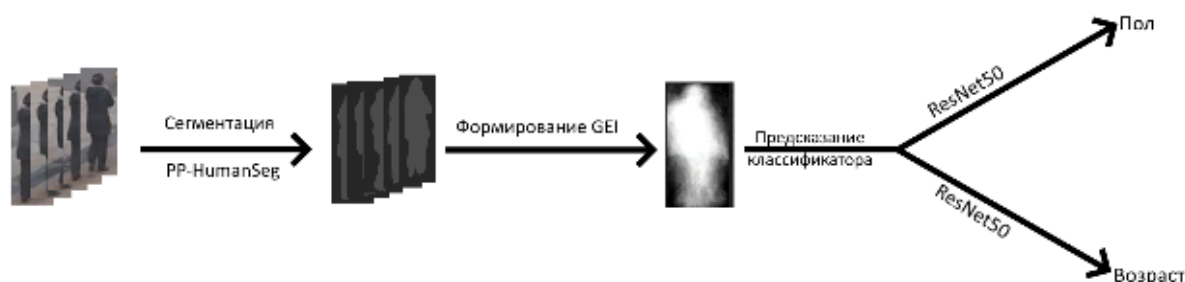


Рис. 1. Конвейер преобразований данных для одного треклета объекта

В таблице 1 указаны результаты экспериментов при разном количестве выходов и разных размерах треклетов.

Таблица 1

Результаты экспериментов по отбору модели классификации пола и возраста

№	Кол-во вых-ов	Размер треклета	Precision				Recall				F1			
			пол	<15	15-45	>45	пол	<15	15-45	>45	пол	<15	15-45	>45
1	1	не огр.	0.75	0.40	0.50	0.63	0.65	0.48	0.22	0.35	0.69	0.43	0.40	0.50
2	1	6	0.76	0.46	0.50	0.65	0.70	0.53	0.25	0.35	0.72	0.51	0.46	0.51
3	2	не огр.	0.81	0.35	0.51	0.44	0.76	0.27	0.22	0.54	0.78	0.36	0.37	0.51
4	2	6	0.83	0.62	0.55	0.48	0.85	0.38	0.28	0.44	0.84	0.52	0.46	0.50

По результатам, приведённым в таблице 6, видно, что более высокое качество как в задаче классификации пола, так и в задаче классификации возраста показывает модель с двумя выходами. С точки зрения размера треклета, лучшее качество показывают модели, для обучения которых использовались треклеты фиксированного размера. В результате, наилучшей моделью, показывающей метрики качества классификации пола 0.84 по мере F1 и возраста в среднем 0.49 по мере F1, является нейронная сеть с двумя выходами для одновременной классификации пола и возраста, использующая GEI-изображения, построенные по треклетам объектов фиксированного размера 6.

Наилучшим методом агрегации множества решений модели является выбор наиболее часто встречающегося класса. Это позволяет улучшить итоговые значения метрик для распознавания пола – до 0,89 по мере F1, для распознавания возраста – до 0,74 по мере F1.

Заключение

В результате проделанной работы был реализован алгоритм для классификации пола и возраста людей в видеопотоке, который использует признаки силуэта и движения человека. Качество алгоритма оценивается 0,89 для распознавания пола и 0,74 для распознавания возраста по мере F1.

Список использованных источников

1. Zhao J. Look Across Elapse: Disentangled Representation Learning and Photorealistic Cross-Age Face Synthesis for Age-Invariant Face Recognition / J. Zhao, Y. Cheng // Proceedings of the AAAI conference on artificial intelligence. 2019. – Vol. 33 № 01. – P. 9251-9258.
2. Lassner C. Unite the People: Closing the Loop Between 3D and 2D Human Representations / C. Lassner, J. Romero, M. Kiefel [at al.] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – P. 4704-4713.
3. Deng, Y. Pedestrian Attribute Recognition At Far Distance / Y. Deng, P. Luo, C.C. Loy [at al.] // Proceedings of the 22nd ACM international conference on Multimedia. – 2014. – P. 789–792
4. Köstinger M. Large scale metric learning from equivalence constraints / M. Köstinger, M. Hirzer [at al.] // 2012 IEEE Conference on Computer Vision and Pattern Recognition. – 2012 – P. 2288-2295