

ОРГАНИЗАТОРЫ:



ИНСТИТУТ
КОСМИЧЕСКИХ ИССЛЕДОВАНИЙ
РОССИЙСКОЙ АКАДЕМИИ НАУК



ГОСУДАРСТВЕННЫЙ
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ
ИНСТИТУТ АВИАЦИОННЫХ СИСТЕМ



ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
ИМ. М. В. КЕЛДЫША
РОССИЙСКОЙ АКАДЕМИИ НАУК



ВСЕРОССИЙСКАЯ НАУЧНО-ТЕХНИЧЕСКАЯ КОНФЕРЕНЦИЯ

ТЕХНИЧЕСКОЕ ЗРЕНИЕ В СИСТЕМАХ УПРАВЛЕНИЯ 2020

17-18
марта
ТЕЗИСЫ

TVCS^{'20}



УДК [004.896:681.5](063)
ББК 32.81я431(2Рос)

Т38 Техническое зрение в системах управления — 2020: сб. тез. докл. научно-техн. конф. 17–18 марта 2019, Москва, ИКИ РАН. М.: ИКИ РАН, 2020. 44 с.

ISBN 978-5-00015-046-7

Всероссийская научно-техническая конференция «Техническое зрение в системах управления—2020» (ТЗСУ-2020) проводится 17–18 марта 2020 г. в Москве, в Институте космических исследований РАН. Организаторами конференции являются Институт космических исследований РАН (ИКИ РАН), Государственный научно-исследовательский институт авиационных систем (ГосНИИАС) и Институт прикладной математики им. М.В. Келдыша РАН (ИПМ им. М. В. Келдыша РАН).

Целью проведения конференции ТЗСУ-2020, как и всех предыдущих конференций ТЗСУ, являются обмен информацией между академической и прикладной наукой, образованием и производством, совместное обсуждение достигнутых научных и практических результатов, а также обсуждение вопросов, являющихся ключевыми для развития технического зрения и его использования в системах управления.

Тематика конференции охватывает различные аспекты разработки и построения систем технического зрения (СТЗ) в системах управления. Особое внимание уделяется бортовым СТЗ, входящим в состав мобильных объектов и предназначенным для решения задач автономного и автоматизированного управления в сложной, неопределённой и быстро изменяющейся внешней обстановке.

Труды конференции публикуются в рецензируемом электронном журнале «Техническое зрение»: <http://magazine.technicalvision.ru/>.

Для проведения исследования собрана база данных по городу Нью-Йорку, состоящая из карт высот и юстированных спутниковых изображений разрешением 1 м. Карты высот сформированы с использованием лидарной съёмки 2017 г. из общедоступной базы National Lidar Dataset (United States), а спутниковые изображения получены из серверов google, yandex, nokia и bing с использованием QGIS.

Проведена проверка работы предлагаемого алгоритма на тестовой выборке для города Нью-Йорк и на базе данных ISPRS для города Потсдам, средняя квадратическая ошибка которых составила 3,84 и 5,2 м соответственно.

ЛИТЕРАТУРА

1. *Ronneberger O., Fischer Ph., Brox Th.* U-Net: Convolutional Networks for Biomedical Image. 2015.
2. *Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, Hartwig Adam.* SegmentationEncoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. 2018.
3. *Ke Sun, Yang Zhao, Borui Jiang, Tianheng Cheng, Bin Xiao, Dong Liu, Yadong Mu, Xinggang Wang, Wenyu Liu, Jingdong Wang.* High-Resolution Representations for Labeling Pixels and Regions. 2019.

ОБНАРУЖЕНИЕ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ БОЛЬШОГО РАЗРЕШЕНИЯ НА ОСНОВЕ ИХ МНОГОМАСШТАБНОГО ПРЕДСТАВЛЕНИЯ И БЛОЧНОЙ ОБРАБОТКИ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТЬЮ

Р. П. Бозуш¹, И. Ю. Захарова¹, С. В. Абламейко²

¹ Полоцкий государственный университет (ПГУ), Полоцк, Республика Беларусь

² Белорусский государственный университет (БГУ), Минск, Республика Беларусь
bogushr@mail.ru, i.zakharova@psu.by, ablaimeyko@bsu.by

В настоящее время и в ближайшем будущем все шире будут использоваться устройства, позволяющие формировать изображения высокой чёткости больших размеров с разрешением 4К и больше. В этом случае значительно увеличивается качество отображения объектов небольших размеров, либо удалённых на значительное расстояние от видеокамеры при съёмке. Соответственно, предоставляется возможность автоматического обнаружения таких объектов, которые при меньшем разрешении не обладают достаточной информативностью для их идентификации.

В последнее время благодаря стремительному развитию вычислительной мощности компьютерной техники, все шире используется подход для обнаружения и идентификации объектов, который основан на применении свёрточных нейронных сетей (СНС). Однако на первом этапе при использовании СНС выполняется уменьшение размера исходного изображения к размерам входного слоя, что ведёт к понижению информативности объектов, а при их небольших размерах это приводит к невозможности их обнаружения на изображении после его масштабирования. В качестве СНС используется модель YOLO третьей версии, так как она характеризуется меньшими вычислительными затратами и высокой точностью. Определение количества уровней пирамиды изображений осуществляется с учётом того, что размеры самого верхнего уровня должны быть приближены к размерам изображения входного слоя используемой СНС. В результате экспериментов установлено, что для обеспечения высокой точности при обнаружении объектов небольших размеров на изображениях 4К необходимо минимизировать уменьшение размера изображения. Так, при размере входного слоя 608×608 значение метрики mAP составляет 21 %, а при размере 1024×1024 mAP оценивается значением 47 % на тестовом наборе изображений с разрешением 4К при пороговой степени уверенности 0,5. Таким образом, результативность обнаружения и идентификации объектов для изображения верхнего уровня определяется характеристиками СНС. Для других масштабов выполняется разбиение изображений на перекрывающиеся блоки, что позволит улучшить правильную классификацию объектов, фрагменты которых будут отнесены к соседним блокам.

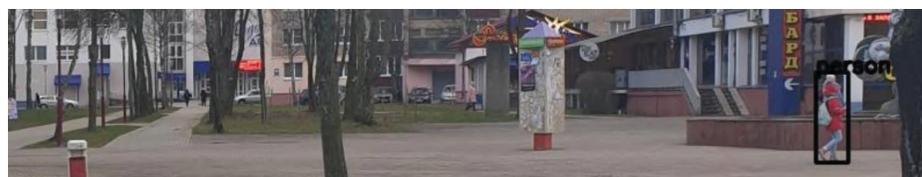
Обнаруженная с использованием СНС область описывается набором таких признаков, как: координаты верхнего левого и нижнего правого углов найденной области интереса; класс объекта; уверенность СНС в правильной классификации объекта; номер уровня пирамиды на которой объект был найден. На разных масштабах изображений возможно незначительное отличие координат локализации объекта. Кроме этого, фрагмент объекта большого размера может быть идентифицирован как независимая область-кандидат на одном из средних или нижних уровней. Поэтому решение об объединении

граничных объектов соседних блоков принимается на основе анализа величины их взаимного перекрытия и принадлежности к одному классу.

Результаты выполненных исследований показали, что на исследуемой базе данных 4К-изображений, содержащих большое количество объектов малых размеров, значение метрики mAP для разработанного алгоритма при размере входного слоя СНС 1024×1024 составляет 75,9 %, т.е. значительно выше, чем при использовании СНС YOLO v3. На рис. 1 показаны фрагменты обработанных 4К-изображений с обнаруженными объектами, людьми и машинами. С использованием предложенного алгоритма идентифицировано 9 объектов (рис. 1а), минимальный размер из них 33×13 , а применение СНС YOLO v3 позволило обнаружить один объект размером 149×54 (рис. 1б). Таким образом, предложенный подход является перспективным для детектирования объектов на изображениях большого разрешения.



а



б

Рис. 1. Фрагменты 4К-изображения с примерами обнаружения объектов:
а — для предложенного алгоритма; б — для СНС YOLO v3

ИНТЕГРАЦИЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ КЛАССИФИКАЦИИ ОБЪЕКТОВ В СИСТЕМУ ИНТЕЛЛЕКТУАЛЬНОГО ВИДЕОНАБЛЮДЕНИЯ

И. С. Фомин¹, В. В. Бурин², А. В. Бахшиев^{1,2}

¹ Центральный научно-исследовательский институт робототехники и технической кибернетики, Москва, Россия, i.fomin@rtc.ru

² Санкт-Петербургский политехнический университет Петра Великого (СПбПУ) Санкт-Петербург, Россия, bvv98@yandex.ru, bakhshiev_av@spbstu.ru

Статья посвящена проблемам интеграции свёрточных нейронных сетей в состав системы интеллектуального видеонаблюдения. На сегодняшний день существует большое количество архитектур нейронных сетей, позволяющих классифицировать изображения. Серьёзное развитие этой области началось с появлением в свободном доступе по доступной цене графических процессоров (GPU), вычислительные способности которых исследователи смогли применить для матричных и векторных вычислений, лежащих в основе нейронных сетей. Начиная с первой известной работы (Krizhevsky et al., 2012) развитие области нейронных сетей было очень активным. Вскоре появились решения с существенно увеличенной глубиной (количеством слоёв) VGG-16 и VGG-19 (Simonyan, Zisserman, 2014), поднявшие качество работы сетей на новый уровень. Когда увеличение числа слоёв перестало приносить желаемый результат, так как появилась проблема исчезающих градиентов, и метод градиентного спуска, обычно используемый для обучения слоёв, перестал оказывать на веса первых слоёв заметное влияние и обучение перестало сходиться. После этого исследователи нашли способы решения этой проблемы. В сети GoogLeNet (Inception V1) (Szegedy et al., 2015) впервые были предложены примеры модулей — строительных блоков из которых собирается сеть, в которых присутствует несколько слоёв, соединённых параллельно так, что данные при обучении могут идти по разным путям.