

Нижегородский государственный университет им. Н.И. Лобачевского
Институт информационных технологий, математики и механики
Кафедра Математического обеспечения и суперкомпьютерных технологий

Образовательный курс
«Современные методы и технологии глубокого обучения
в компьютерном зрении»

Лекция № 5
Глубокие модели для сопровождения объектов на видео

При поддержке компании Intel

Васильев Е.П.

Нижний Новгород
2020

Содержание

1	Аннотация	3
2	Литература	3
2.1	Основная литература	3
2.2	Ресурсы сети Интернет	4

1 Аннотация

Цель данной лекции состоит в том, чтобы изучить глубокие нейросетевые модели для решения *задачи сопровождения объектов на видео* (object tracking).

В начале лекции дается постановка задачи сопровождения объектов. Затем приводится обзор широко известных наборов данных и бенчмарков для сопровождения объектов: Multiple Object Tracking Challenge [1, 10], Long-Term Visual Object Tracking Benchmark [2, 11], TrackingNet [3, 12] и другие. Описываются наиболее используемые показатели качества сопровождения объектов. Отдельно рассматриваются показатели сопровождения одиночных объектов: точность (accuracy) и надежность (robustness). Определяются показатели сопровождения нескольких объектов: достоверность (Multiple Object Tracking Accuracy, MOTA) и точность сопровождения нескольких объектов (Multiple Object Tracking Precision, MOTP) [4]. Далее приводится классификация методов сопровождения по доступу к данным и классификация глубоких моделей по способу их использования в схеме сопровождения через сопоставление объектов. Дается обзор глубоких моделей для решения задачи сопровождения объектов. Вначале описывается модель DeepSORT [5], которая обеспечивает построение описания изображения объекта с использованием глубоких моделей для последующего сопровождения. Далее рассматриваются методы SINT [6], SiameseNET [7] и GOTURN [8], которые предполагают встраивание глубоких нейросетевых моделей в разные этапы алгоритма сопровождения. В заключении описывается модель, построенная на базе рекуррентных сетей, RNN-LSTM [9], которая полностью решает задачу сопровождения.

В настоящее время происходит развитие алгоритмов сопровождения объектов на видео, появляются все более крупные наборы данных и разрабатываются новые глубокие модели. На текущий момент задачу сопровождения объектов нельзя считать полностью решенной. Однако, программные системы, способные сопровождать конкретные классы объектов (автомобили, пешеходы) с высокой точностью уже существуют и применяются в промышленности, производстве, ритейле.

2 Литература

2.1 Основная литература

1. Leal-Taixé L., Milan A., Reid I., Roth S., Schindler K. MOTChallenge 2015: Towards a Benchmark for Multi-Target Tracking. – 2015. – [<https://arxiv.org/pdf/1504.01942.pdf>].
2. Moudgil A., Gandhi V. Long-Term Visual Object Tracking Benchmark. – 2019. – [<https://arxiv.org/abs/1712.01358>].
3. Müller M., Bibi, A., Giancola S., Al-Subaihi S., Ghanem B. TrackingNet: A Large-Scale Dataset and Benchmark for Object Tracking in the Wild. – 2018. – [<https://arxiv.org/abs/1803.10794>].
4. Bernardin K., Stiefelhagen R. Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics // Image and Video Processing. – 2008.
5. Wojke N., Bewley A., Paulus D. Simple online and realltime tracking with a deep association metric // International Conference on Image Processing. – 2017. – P. 3645–3649. – [<https://arxiv.org/pdf/1703.07402.pdf>], [<https://ieeexplore.ieee.org/document/8296962>].
6. Tao R., Gavves E., Smeulders A. Siamese instance search for tracking. – 2016. – [<https://arxiv.org/pdf/1605.05863.pdf>], [<https://ieeexplore.ieee.org/document/7780527>].
7. Leal-Taixé L., Canton-Ferrer C. and Schindler. K. Learning by tracking: siamese CNN for robust target association. – 2016. – [<https://arxiv.org/pdf/1604.07866.pdf>], [<https://ieeexplore.ieee.org/document/7789549>].
8. Held D., Thrun S., Savarese S. Learning to track at 100 FPS with deep regression networks. – 2016. – [<https://arxiv.org/pdf/1604.01802.pdf>].
9. Milan A., Rezatofighi S.H., Dick A.R., et al. Online multi-target tracking using recurrent neural networks. – 2017. – [<https://arxiv.org/pdf/1604.03635.pdf>], [<https://dl.acm.org/doi/10.5555/3298023.3298181>].

2.2 Ресурсы сети Интернет

10. Multiple Object Tracking Benchmark [<https://motchallenge.net>].
11. Long-Term Visual Object Tracking Benchmark [<https://amoudgl.github.io/ltot>].
12. TrackingNet [<https://tracking-net.org>].