Transferring Rich Feature Hierarchies for Robust Visual Tracking

mingzailao

2016-09-08 Thu

Outline

Introduction

Goal

Given this single (labeled) instance, the goal is to track the movement of the object in an online manner.

The Contributions

- To alleviate the overfitting and drifting problems during online tracking, we pre-train the CNN to distinguish objects from non-objects instead of simply reconstructing the input or performing categorical classification on large-scale datasets with object-level annotations.

Introduction

The Contributions

- The output of the CNN is a pixel-wise map to indicate the probability that each pixel in the input image belongs to the bounding box of an object. The key advantages of the pixel-wise output are its induced structured loss and computational scalability.
- We evaluate our proposed method on an open benchmark as well as a challenging non-rigid object tracking dataset and obtain very remarkable results. In particular, we improve the area under curve (AUC) metric of the overlap rate curve from 0.529 to 0.602 for the open benchmark.

Overview

Two Stages

- offline pre-training stage
- online fine-tuning and tracking stage

Pre-training Stage

We train a CNN to learn generic object features for distinguishing objects from non-objects.

Tracking Stage

Fine-tuning the parameters so that CNN can adapt to the target.

Overview

For Robustness

Running two CNNs concurrently during online tracking to account for possible mistakes caused by model update.