Sports Object Recognition and Tracking⋆

Shreyansh Sharma[1][S3772241]

Leiden University, Netherlands
S3772241@vuw.leidenuniv.nl

Abstract. Sports object recognition and tracking is a challenging task in computer vision. This paper presents and compares different methods for object recognition and tracking in sports videos. Th focus lies on utilizing pre-trained models to reduce the training time and improve the accuracy of the model. Along with the comparison, a framework is proposed to evaluate the performance of the models.

Keywords: DETR · DeepSort · SportsMot.

1 Introduction

In recent years, the field of computer vision has seen a lot of progress in object recognition and tracking. Improved object recognition and tracking can be used in many applications such as autonomous driving, surveillance, and sports analysis. In sports, object recognition and tracking can be used to analyze the performance of the players and the team. It can also be used to analyze the performance of the referee and the umpire. The data generated from object recognition and tracking can be used for downstream tasks such as player tracking, player action recognition, and player pose estimation.

Object recognition and tracking in sports is a challenging task due to the fast movement of the players and the ball. The players and the ball can be occluded by other players or the referee. The players can also be occluded by the audience. This makes it difficult to track the players and the ball.

In this paper, we present and compare different methods for object recognition and tracking in sports videos. The focus lies on utilizing pre-trained models to reduce the training time and improve the accuracy of the model. Along with the comparison, a framework is proposed to allow faster setup and evaluation of the models.

For the comparison, we use the SportsMot dataset [1]. The SportsMot dataset is a large-scale dataset for multi-object tracking in sports. It contains videos of different sports such as basketball, football, and volleyball. This makes it suitable for comparing different methods for object recognition and tracking in sports. MOT-16 and MOT-17 datasets [2] are also popular datasets for multi-object tracking.

⋆ Supported by organization Leiden University.

2   Related Work

Usually, the task of object recognition and tracking is tackled in two steps. First, the objects are detected in each frame of the video. Second, the detected objects are tracked across the frames of the video. This is also commonly known as the tracking-by-detection approach.

In recent years, there has also been a lot of progress end-to-end object recognition and tracking.

2.1   Object Detection

2.2   Object Tracking

2.3   A Subsection Sample

Please note that the first paragraph of a section or subsection is not indented. The first paragraph that follows a table, figure, equation etc. does not need an indent, either.

Subsequent paragraphs, however, are indented.

Sample Heading (Third Level) Only two levels of headings should be numbered. Lower level headings remain unnumbered; they are formatted as run-in headings.

Sample Heading (Fourth Level) The contribution should contain no more than four levels of headings. Table 1 gives a summary of all heading levels.

**Table 1.** Table captions should be placed above the tables.

Heading level	Example	Font size and style
Title (centered)	**Lecture Notes**	14 point, bold
1st-level heading	**1 Introduction**	12 point, bold
2nd-level heading	**2.1 Printing Area**	10 point, bold
3rd-level heading	**Run-in Heading in Bold.** Text follows	10 point, bold
4th-level heading	*Lowest Level Heading.* Text follows	10 point, italic

Displayed equations are centered and set on a separate line.

$$x + y = z \tag{1}$$

Please try to avoid rasterized images for line-art diagrams and schemas. Whenever possible, use vector graphics instead (see Fig. 1).

Theorem 1. *This is a sample theorem. The run-in heading is set in bold, while the following text appears in italics. Definitions, lemmas, propositions, and corollaries are styled the same way.*
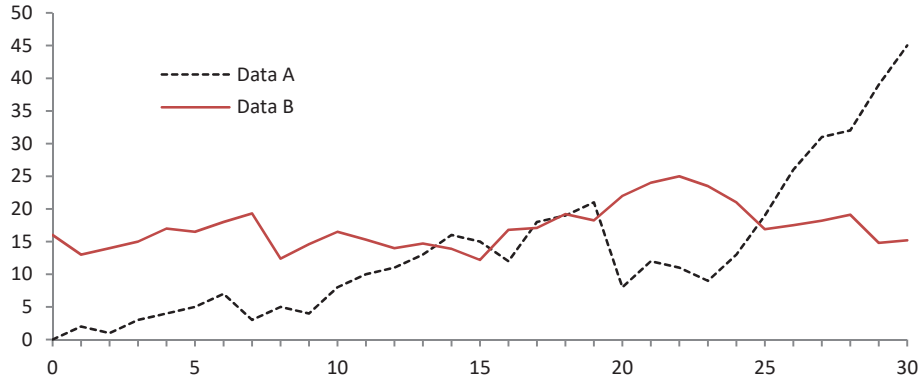
Fig. 1. A figure caption is always placed below the illustration. Please note that short captions are centered, while long ones are justified by the macro package automatically.

Proof. Proofs, examples, and remarks have the initial word in italics, while the following text appears in normal font.

For citations of references, we prefer the use of square brackets and consecutive numbers. Citations using labels or the author/year convention are also acceptable. The following bibliography provides a sample reference list with entries for journal

References

1. Cui, Y., Zeng, C., Zhao, X., Yang, Y., Wu, G., Wang, L.: Sportsmot: A large multi-object tracking dataset in multiple sports scenes. arXiv preprint arXiv:2304.05170 (2023)
2. Milan, A., Leal-Taixe, L., Reid, I., Roth, S., Schindler, K.: Mot16: A benchmark for multi-object tracking (2016)