Enhancing Player Tracking Across Sports Domains Through Meta-Learning with Video Input

Stanford University
CS 330: Deep Multi-Task and Meta Learning, Fall 2023

Akayla Hackson
akayla@stanford.edu

Lavinia Pedrollo
laviniap@stanford.edu

Miguel Gerena
miguelg2@stanford.edu

1 Objective

Our project's primary objective is to leverage meta-learning approaches for the effective utilization of video input, enhancing player tracking across diverse sports domains. In an era marked by rapid advancements in machine learning, computer vision, and data analysis, exploring uncharted territories becomes imperative. By introducing meta-learning to sports player tracking through video analysis in different sport domains, we push the boundaries of what is currently possible, positioning ourselves as pioneers in the field.

The concept of tracking players across different sports domains is highly relevant in our interconnected world. Adapting and generalizing tracking techniques to various sports is paramount, aligning with the evolving landscape of interdisciplinary research. Our objective to employ meta-learning approaches for tracking players across diverse sports domains addresses a crucial need. This adaptability breaks down silos between different sports, fostering cross-disciplinary insights that can drive innovation in sports analytics. The limited baselines for utilizing meta-learning in sports video analysis make our project even more relevant. By addressing this gap, we aim to establish a foundation for future research and development.

Accurate player tracking is invaluable for teams, coaches, and analysts, aiding in understanding player movements, making data-driven decisions, and enhancing overall sports performance. Our project has the potential to significantly contribute to improving the accuracy and efficiency of player tracking, transforming the way sports are analyzed and coached. This relevance positions our work at the forefront of advancements in the sports industry and resonates with enthusiasts seeking cutting-edge solutions.

2 Related Work

The field of player detection and tracking in sports videos is rapidly evolving thanks to the latest developments in machine learning and artificial intelligence, particularly the growing popularity of computer vision. While there is already a wealth of research on object detection, our project aims to take this a step further. Specifically, we are focused on player detection in sports games. This requires the implementation of multiple object tracking (MOT) throughout the entire video sequence.

Our extensive research in this domain has revealed several existing models, including You Only Look Once (YOLO [1] and its variations like YOLOX[2]), Detectron2[3], TrackNet[4], ByteTrack[5], OC-SORT[6], MixSort-OC[7], and the state-of-the-art (SOTA) Deep-EIoU[8]. These models encompass a wide spectrum of deep learning techniques, ranging from single-shot full convolutional neural networks to heatmap neural networks and various approaches in between.

Given the diversity of these deep learning techniques, it remains uncertain which model will outperform the others when applied to our dataset. Therefore, our plan is to implement and evaluate all these models to determine which one exhibits the best performance for our specific dataset.

3 Technical Outline

The team will conduct further research on state-of-the-art algorithms, dataset selection, and state-of-the-art meta-learning approaches. Currently, the team has identified the SportsMOT Dataset[9] which consists of 240 clips from three different domains (basketball, soccer, and volleyball). All the videos are recorded and downloaded with 720P (standard high) resolution, 25 frames per second (FPS) with all of the clips manually cut to an average of 485 frames each. We also like this dataset for its diversity. The different sport domains have different camera views from one another, as well as different scenery (i.e. Soccer is filmed on grass in an outdoor stadium while basketball is filmed indoors on various wood courts).

The team will focus on the development of an algorithm that demonstrates the meta-learning techniques' capability of adapting to different sports domains, breaking the silos of specialized tracking systems. By leveraging meta-learning techniques, our model is expected to achieve high accuracy and robustness in player tracking across various sports. This is relevant for providing valuable sport analytics insights.

A data pipeline will be created to ensure the same training and test data sets are being utilized on all models and an accurate comparison is made between them. The models developed by the team will be evaluated against the SportsMOT dataset[9] and compared to their SOTA models. In addition, the team will use other datasets as necessary for training and/or evaluation.

The team will then triage on the evaluation of at least three current SOTA traditional models (specialized player tracking systems) and investigate other valid architectures. From this initial investigation, the team will downselect architectures and methodologies, train the candidates, and compare to fine-tuning existing models to improve performance.

The team's main contribution to the field will be increasing the performance of the current SOTA for multi-sport player tracking utilizing non-static datasets (video inputs). In addition, we will be exploring the categorization of each player into a team.

It is important to understand the performance improvements of meta-learning approaches versus traditional methods. The team will be comparing and benchmarking against traditional SOTA methods. By comparing our meta-learning approach to traditional, specialized player tracking methods, we aim to provide a concrete evaluation of the effectiveness of meta-learning approaches.

The models generated will be evaluated utilizing High Order Tracking Accuracy (HOTA) [10], Multi-Object Tracking Accuracy (MOTA) [11], and the ratio of correctly identified detections over the average number of ground-truth and computed detections (IDF1) [12]. These metrics are subject to change as we might add or remove metrics while we progress throughout the project.

(Optional stretch goal) Real-time tracking. If successful, our algorithm could be used in real-time for tracking players on live streams.

4 Team Contributions

- Akayla Hackson will research traditional existing algorithms in use for specialized player tracking systems. Akayla will be responsible for meticulously selecting three algorithms to be later implemented by each of the project's team members for benchmarking.
- Lavinia Pedrollo will collaborate with other team members to address data-related issues and collaboratively conduct data preprocessing efforts, including cleaning, standardization, and augmentation.
- Miguel Gerena will ensure that all the project's team members are up to date on the best performing meta-learning approaches, tailored to our solution.
- Each team member will contribute to the implementation of at least one traditional existing player tracking algorithm used in specialized systems. These traditional algorithms will serve as a benchmark for assessing the adaptability and performance of the new algorithm.

- Each team member will collaborate together to develop a new meta-learning algorithm that will be superior in terms of adaptability and performance when compared to the traditional algorithms.
- Each team member will work on implementing, fine-tuning, and experimenting with different variations of the meta-learning model.
- In the evaluation phase, each team member will test at least one traditional player tracking algorithm and compare the performance with our new meta-learning algorithm.

References

[1] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection, 2016.

[2] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun. Yolox: Exceeding yolo series in 2021, 2021.

[3] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. `https://github.com/facebookresearch/detectron2`, 2019.

[4] Yu-Chuan Huang, I-No Liao, Ching-Hsuan Chen, Tsì-Uí İk, and Wen-Chih Peng. Tracknet: A deep learning network for tracking high-speed and tiny objects in sports applications, 2019.

[5] Yifu Zhang, Peize Sun, Yi Jiang, Dongdong Yu, Fucheng Weng, Zehuan Yuan, Ping Luo, Wenyu Liu, and Xinggang Wang. Bytetrack: Multi-object tracking by associating every detection box, 2022.

[6] Jinkun Cao, Jiangmiao Pang, Xinshuo Weng, Rawal Khirodkar, and Kris Kitani. Observation-centric sort: Rethinking sort for robust multi-object tracking, 2023.

[7] Yutao Cui, Chenkai Zeng, Xiaoyu Zhao, Yichun Yang, Gangshan Wu, and Limin Wang. Sportsmot: A large multi-object tracking dataset in multiple sports scenes, 2023.

[8] Hsiang-Wei Huang, Cheng-Yen Yang, Jiacheng Sun, Jenq-Neng Hwang, and Chung-I Huang. Iterative scale-up expansioniou and deep features association for multi-object tracking in sports, 2023.

[9] Yutao Cui, Chenkai Zeng, Xiaoyu Zhao, Yichun Yang, Gangshan Wu, and Limin Wang. Sportsmot: A large multi-object tracking dataset in multiple sports scenes, 2023.

[10] Jonathon Luiten, Aljosa Osep, Patrick Dendorfer, Philip Torr, Andreas Geiger, Laura Leal-Taixe, and Bastian Leibe. Hota: A higher order metric for evaluating multi-object tracking. *International Journal of Computer Vision*, 2020.

[11] Saif Hassan, Ghulam Mujtaba, Asif Rajput, and Noureen Fatima. Multi-object tracking: a systematic literature review. *Multimedia Tools and Applications*, 2023.

[12] M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus. Performance measures and a data set for multi-target, multi-camera tracking. 2016.