**Soccernet-Player Tracking Challenge Proposal**

| **Aditya Bothra** | **Tyler Banks** |
| --- | --- |
| **MS CSE, University At Buffalo** | **MS AI, University At Buffalo** |
| **abothra@buffalo.edu** | **tylerban@buffalo.edu** |

* **What is the problem?**

The task chosen for this challenge is Multi-Player Tracking. Multi-object tracking (MOT) aims at recovering trajectories of multiple objects in time by estimating object bounding boxes and identities in video sequences. We consider two tasks: (1) a pure association task that considers ground-truth detections (the task of the first challenge), or (2) a complete tracking task that expects to detect the objects of interest from the raw video (a task for benchmarking and future challenges).

* **Motivation**

One motivation for this challenge is to aid in building an advanced automatic player tracking solution. With successful tracking of multiple players, we should be able to use this information to easily see all the frames where a particular player impacts the game or we can segment highlights payer-wise. This can be used for customized videos that focus on a single-player or a premium spotlight star view with a personalized display of statistics side by side.

* **Related work in the research literature (ongoing):**

MOT.MOT17, MOT18 versions of multiple objects Tracking

Deep Sort, ByteTrack model for multiple object detection and tracking,

Association, and detection using Computer Vision.

Baidu’s Baseline repository for soccer-net 2021 [SoccerNetBenchmarks](https://github.com/SoccerNet/sn-tracking/tree/main/Benchmarks)

* **Data set is planned to be used**

The Tracking dataset consists of 12 complete soccer games from the main camera including

* 200 clips of 30 seconds with tracking data.
* one complete halftime annotated with tracking data.
* the complete videos for the 12 games.

Note that a subset of this data is used in this first challenge. In particular, this accounts for 57 30-seconds clips for the train set, 49 clips for the test set, 58 clips for our first public challenge, and 37 clips for our future challenges, including the entire half-time video in the latter.

* **Compute resources needed**

Google Colab Pro for GPU for background execution.

Google Drive Storage for 200 GB.

* **Performance metrics**

The object classes are not taken into account in this challenge or the evaluation. The objects to retrieve are among the following classes: players, goalkeepers, referees, balls, and any other human entering the field.

For the benchmark and challenge, Baidu considers HOTA as the main metric. More specifically, this metric can be decomposed into two components: DetA and AssA, focusing on detection and association accuracy, respectively.

* **Team members and expected roles**

| **Team Member** | **Task** |
| --- | --- |
| **Tyler Banks** | **Image and Video processing for segmentation, Tracking AssA** |
| **Aditya Bothra** | **Occlusion, Architecture Design for Deep Learning, DetA** |

* **References**

[**SoccerNet - Re-Identification (soccer-net.org)**](https://www.soccer-net.org/tasks/re-identification)

**https://github.com/SoccerNet/sn-tracking**