Football Player Tracking with YOLO, ByteTrack, and OSNet

Mohd Zubair Khan

July 11, 2025

Introduction

This project focuses on tracking football players, referees, and the ball in match videos. The main goal is to assign consistent IDs to each player, even after occlusion or re-entry, using a combination of object detection, tracking, and re-identification techniques.

Approach and Methodology

1. **Object Detection:**
   I used the YOLO (You Only Look Once) model to detect players, referees, and the ball in each video frame. YOLO is known for its speed and accuracy in real-time object detection tasks.

2. **Object Tracking:**
   For tracking detected objects across frames, I implemented the ByteTrack algorithm. ByteTrack helps maintain object identities as they move, even in crowded scenes.

3. **Player Re-Identification (ReID):**
   To ensure that each player keeps the same ID even after leaving and re-entering the frame, I explored two main techniques:

   - **OSNet (Deep Feature Extraction):** I used OSNet, a deep learning model for person re-identification, to extract appearance features from each player crop.
   - **Color Histogram:** I computed color histograms (in HSV space) for each player, focusing on jersey color as a distinguishing feature.

   The final ID assignment combines both color histogram similarity and OSNet feature similarity, along with spatial proximity, for robust matching.

Techniques Tried and Outcomes

- **OSNet Only:**
  Initially, I relied solely on OSNet for player re-identification. However, the results

were not satisfactory, especially when players had similar appearances or the model failed to generalize to football jerseys.

- **Color Histogram Only:**
  I then experimented with using only color histograms. This approach yielded much better results, as jersey color is a strong distinguishing feature in football.
- **Combined Approach:**
  Finally, I combined both OSNet features and color histograms. I used color histogram similarity as a fast pre-filter, followed by OSNet feature comparison for fine-grained matching. This hybrid approach provided the most robust and consistent player IDs.

Challenges Encountered

- **OSNet Generalization:**
  One major challenge was that OSNet, trained on generic person re-identification datasets, did not perform well on football players in jerseys. The feature vectors were not always discriminative enough, leading to ID switches.
- **Color Similarity:**
  While color histograms worked well for most cases, they struggled when teams had similar jersey colors or lighting conditions changed.
- **Implementation of OSNet:**
  Integrating OSNet required understanding its API and adapting it for single-image feature extraction. I referred to the official OSNet GitHub repository (`https://github.com/KaiyangZhou/deep-person-reid`) for implementation details and troubleshooting.
- **Balancing Speed and Accuracy:**
  Ensuring real-time performance while maintaining high accuracy was challenging, especially when combining multiple feature extraction methods.

Conclusion

This project demonstrates a robust approach to football player tracking by combining state-of-the-art detection, tracking, and re-identification techniques. The hybrid use of color histograms and deep features significantly improved ID consistency. Future work could focus on team classification, player statistics, and event detection.

References

- Ultralytics YOLO: `https://github.com/ultralytics/ultralytics`
- ByteTrack: `https://github.com/ifzhang/ByteTrack`
- OSNet / torchreid: `https://github.com/KaiyangZhou/deep-person-reid`
- Supervision: `https://github.com/roboflow/supervision`