Deep Learning

Project phase 1
Object Tracking

Dr.Fatemizadeh

Sharif University of Technology

Electrical Engineering

Tina Halimi _ 400101078

Date: February 4, 2025

Contents

1 Algorithms Explanations

In this section, we provide a detailed review of four multi-object tracking algorithms: SORT, DeepSORT, FairMOT, and ByteTrack. Each algorithm is analyzed in terms of its core techniques, strengths, weaknesses, and overall contributions to the field.

1. SORT (Simple Online and Realtime Tracking)

Core Idea: SORT is a pragmatic tracking-by-detection method focused on efficiency for online and real-time applications.

Core Techniques:

- **Kalman Filter:** Predicts the future location of each tracklet based on its current state, including bounding box position, aspect ratio, height, and velocity.
- **Hungarian Algorithm:** Matches detections to tracklets using Intersection-over-Union (IoU) as the association metric.
- **Track Handling:** Tracking decisions are made frame-by-frame, associating detections only with the previous frame.

Strengths:

- **Simplicity:** The algorithm is easy to implement and computationally lightweight.
- **Speed:** Extremely fast, making it suitable for real-time applications.

Weaknesses:

- **Identity Switches:** Prone to frequent identity switches during occlusions or crowded scenes.
- **Limited in Complex Scenarios:** Struggles in crowded or dynamic scenes with fast-moving objects.

2. DeepSORT (Deep Association Metric for SORT)

Core Idea: DeepSORT enhances SORT by incorporating deep appearance features, significantly improving tracking robustness and reducing identity switches.

Core Techniques:

- **Kalman Filter:** Predicts motion and updates track states, as in SORT.
- **Deep Appearance Descriptor:** Extracts feature embeddings for each detection using a convolutional neural network (CNN) trained on a person re-identification dataset.
- **Combined Matching Metric:** Combines IoU-based motion matching with appearance feature similarity for robust data association.

Strengths:

- **Reduced Identity Switches:** Effectively combines motion and appearance features to improve tracking.
- **Improved Occlusion Handling:** Can re-identify objects after temporary occlusions.

Weaknesses:

- **Computational Complexity:** Appearance feature extraction adds computational overhead, reducing speed.
- **Reliance on Detection Quality:** Accuracy depends on the performance of the object detector.

3. FairMOT (Fair Multi-Object Tracking)

Core Idea: FairMOT addresses the imbalance between detection and re-identification (Re-ID) tasks by treating them equally, resulting in a fair and unified tracker.

Core Techniques:

- **Anchor-Free Detection:** Built on CenterNet, avoids anchor-based methods to reduce ambiguity and improve fairness.
- **Homogeneous Branches:** Uses separate branches for detection and Re-ID tasks, optimizing them independently.
- **Low-Dimensional Re-ID Features:** Employs 128-dimensional embeddings to balance accuracy and efficiency.
- **Multi-Task Learning:** Balances losses for detection and Re-ID tasks using an uncertainty loss mechanism.

Strengths:

- **High Accuracy:** Achieves state-of-the-art results by balancing detection and Re-ID.
- **Real-Time Performance:** Runs at 30 FPS on high-end GPUs.
- **Robustness:** Handles dense environments effectively with strong occlusion recovery.

Weaknesses:

- **Implementation Complexity:** Challenging to implement due to its detailed design and training requirements.

4. ByteTrack (Multi-Object Tracking by Associating Every Detection Box)

Core Idea: ByteTrack innovatively incorporates low-confidence detections into the tracking process, improving tracking performance, particularly for occluded or hard-to-detect objects.

Core Techniques:

- **Two-Step Matching:**
  - High-confidence detections are matched to tracklets using IoU or appearance similarity.
  - Unmatched tracklets are matched to low-confidence detections using IoU only.
- **Kalman Filter:** Predicts object motion and facilitates matching.
- **Detection Separation:** Divides detections into high- and low-confidence sets based on a confidence threshold.

Strengths:

- **Utilization of Low-Confidence Detections:** Recovers occluded or hard-to-detect objects.
- **State-of-the-Art Performance:** Achieves top results on multiple benchmarks while maintaining real-time speed.
- **Generalizability:** Can be integrated into other trackers for performance enhancement.

Weaknesses:

- **No Re-ID Features:** Relies solely on motion-based matching, limiting performance in crowded scenarios.
- **Detector Dependence:** Accuracy depends heavily on the quality of the object detector.

Comparison of the Algorithms

Metric	**SORT**	**DeepSORT**	**FairMOT**	**ByteTrack**
Accuracy	Moderate	High	Very High	Very High
Speed	Very High	Moderate	Moderate	High
Complexity	Very Low	Moderate	High	Moderate
Re-ID Support	Not Supported	Supported	Fully Integrated	Not Supported
Occlusion Handling	Poor	Good	Excellent	Very Good

Table 1: Comparison of SORT, DeepSORT, FairMOT, and ByteTrack.

Summary: SORT is fast and simple but lacks robustness. DeepSORT improves identity preservation with appearance features. FairMOT achieves state-of-the-art accuracy by balancing detection and Re-ID tasks. ByteTrack excels at occlusion recovery by leveraging low-confidence detections, offering a robust and efficient solution.

2 Detailed Review of Articles and Methods

2.A Review of Articles

2.A.a SoccerNet-Tracking Dataset

Overview: The SoccerNet-Tracking dataset is a benchmark tailored for evaluating tracking algorithms in soccer scenarios. It focuses on addressing key challenges:

- **Visual Similarity:** Players from the same team look nearly identical, making it difficult to distinguish between them.
- **Occlusions:** Frequent overlaps during dense interactions (e.g., corners or penalties) disrupt tracking.
- **Fast-Moving Objects:** The soccer ball's rapid and erratic motion makes it particularly hard to track.

Dataset Features:

- **Composition:** Includes 200 video clips (30 seconds each) and one 45-minute match half. Videos are recorded at 1080p resolution and 25 FPS to support high-quality tracking.
- **Annotations:** Provides bounding boxes for players, referees, and the ball. Includes unique IDs for long-term tracking, even when objects leave and re-enter the frame.
- **Scenarios Captured:** Features complex soccer scenarios like corners, penalties, free kicks, and substitutions, providing a comprehensive testbed for tracking algorithms.

Benchmark Results:

- **ByteTrack:** Achieved the highest HOTA scores, excelling in crowded scenarios and during ball tracking.
- **FairMOT:** Demonstrated strong re-identification performance, particularly for players in dense interactions.
- Fine-tuning on SoccerNet significantly improved ID switch handling and tracking accuracy.

Applications:

1. **Ball Tracking:** ByteTrack's ability to handle low-confidence detections ensures robust ball tracking during high-speed movements or occlusions. Critical for calculating ball possession, pass trajectories, and shot analysis in soccer analytics.
2. **Player Tracking:** FairMOT excels in tracking players during crowded scenarios like corners or free kicks, ensuring accurate re-identification. Essential for analyzing player positioning, formation changes, and tactical strategies.
3. **Referee Tracking:** Provides consistent tracking of referees, useful for evaluating positioning and decision-making during critical events like fouls or offside calls.

2.A.b GTA: Global Tracklet Association for Sports

Overview: GTA is a post-processing algorithm that refines MOT outputs by addressing two critical issues:

- **Fragmented Tracklets:** Tracks split into segments due to occlusions or frame loss.
- **Identity Mix-Ups:** Tracks incorrectly merging data from multiple objects.

Methodology:

- **Tracklet Splitter:** Uses DBSCAN clustering and re-identification (Re-ID) features to separate mixed-identity tracklets.
- **Tracklet Connector:** Employs hierarchical clustering to merge fragmented tracklets using spatial and temporal consistency.

Results:

- Reduced ID switches by 58% and improved HOTA scores by 6.84% on SoccerNet.
- Enhanced tracking robustness during dense interactions.

Applications:

1. **Crowded Scenarios:** Refines outputs from FairMOT or ByteTrack during goal celebrations or corners, ensuring smooth player tracking despite occlusions.
2. **Long-Term Tracking:** Maintains consistent tracks for players or referees who temporarily leave the frame during substitutions or transitions.
3. **Reducing Errors in Basketball and Soccer:** In basketball, where frequent overlaps occur, GTA ensures smoother tracking for downstream tasks like trajectory analysis or player heatmaps.

2.A.c Deep HM-SORT

Overview: Deep HM-SORT enhances SORT by integrating deep learning-based appearance features and advanced motion-appearance association strategies. It addresses challenges like:

- **Appearance Similarity:** Differentiating players in similar jerseys.
- **Unpredictable Movements:** Handling erratic player motions.
- **Occlusions:** Managing overlaps during close player interactions.

Methodology:

- **Deep Features:** Extracts robust appearance embeddings to differentiate players with subtle visual distinctions.
- **Harmonic Mean Association:** Balances appearance and motion cues for more reliable track associations.
- **Expansion IoU:** Enhances overlap calculations to handle occlusions effectively.

Results:

- Achieved state-of-the-art IDF1 and HOTA scores on SoccerNet and SportsMOT, reducing identity switches and improving robustness.

Applications:

1. **Soccer Player Tracking:** Tracks players during fast-paced transitions or counterattacks, ensuring identity consistency in dynamic plays.
2. **Basketball Defensive Analysis:** Tracks players during defensive rotations, even when rapid movements and occlusions occur.
3. **Volleyball Team Rotations:** Tracks players during rallies, maintaining accuracy even when players cluster at the net.

2.A.d SportsTrack

Overview: SportsTrack is a robust MOT framework designed for dynamic sports scenes involving unstable camera views, overlapping players, and motion blur.

Methodology:

- **Three-Stage Matching Process:** Iteratively refines associations between detections and tracks to reduce false matches caused by motion blur or overlapping objects.
- **One-to-Many Correspondence:** Handles player overlaps by associating single detections with multiple tracks when necessary.

Results:

- Outperformed ByteTrack and BOT-SORT on SportsMOT, particularly in challenging basketball and soccer scenarios.

Applications:

1. **Unstable Camera Movement in Soccer:** Tracks players and the ball during rapid camera transitions, ensuring no objects are lost.
2. **Crowded Basketball Scenarios:** Maintains accurate tracking during overlapping plays like pick-and-rolls or screens.
3. **Motion Blur in Volleyball:** Tracks players during fast-paced rallies, mitigating errors caused by motion blur.

2.A.e Basketball-SORT

Overview: Basketball-SORT addresses the specific challenges of complex multi-object occlusions (CMOOs) in basketball. It combines motion-based predictions and appearance cues to ensure accurate tracking.

Methodology:

- **Trajectory-Based Association:** Predicts player movements to maintain tracking during occlusions.
- **Reacquiring Long-Lost IDs (RLLI):** Reassigns identities to players who reappear after leaving the frame.

- **Domain-Specific Constraints:** Introduces court-specific knowledge to improve tracking accuracy.

Results:

- Achieved a HOTA score of 63.48%, demonstrating strong performance in dense scenarios.

Applications:

1. **Basketball Rebounding Analysis:** Tracks players during rebounds, ensuring consistent identity assignment even when players overlap.
2. **Soccer Substitutions:** Maintains player tracks during substitutions or sideline transitions.
3. **Volleyball Tracking:** Tracks all players during dense net interactions, maintaining accuracy despite occlusions.

2.A.f MixSort Framework

Overview: MixSort combines motion-based and appearance-based tracking, addressing the limitations of traditional IoU-only or appearance-only methods.

Methodology:

- **Appearance-Based Association:** Enhances re-identification by integrating robust appearance embeddings.
- **Motion-Based Association:** Uses motion cues to ensure smoother tracking of fast-moving objects.

Results:

- Achieved state-of-the-art performance on SportsMOT and MOT17, demonstrating versatility across sports scenarios.

Applications:

1. **Soccer Free Kicks:** Tracks players during set pieces, ensuring identity consistency despite dense formations.
2. **Volleyball Spikes and Blocks:** Maintains accurate player tracking during close interactions at the net.
3. **Tennis Rallies:** Tracks players during fast-paced rallies, ensuring seamless trajectory analysis.

Summary of Algorithm Applications in Sports Scenarios

ByteTrack, FairMOT, DeepSORT, and SORT play crucial roles in addressing the unique challenges of sports tracking. **ByteTrack** is particularly well-suited for tracking fast-moving objects like the ball due to its ability to handle low-confidence detections and occlusions, making it indispensable for analyzing ball possession and trajectory in soccer or tennis. **FairMOT**, with its integrated detection and re-identification capabilities, excels in tracking players in crowded scenarios, such as goal celebrations or set pieces, providing robust player identity consistency. **DeepSORT** enhances SORT by adding appearance-based features, which makes it effective in handling occlusions and visually similar players, especially in team sports like basketball and volleyball. Finally, **SORT**, though simpler, is an efficient solution for scenarios with minimal occlusion and dynamic motion, suitable for real-time analysis of player trajectories. Together, these algorithms enable comprehensive sports analytics by addressing the challenges of dynamic environments, occlusions, and fast-paced actions.

2.B Analysis of Algorithms in Sports Scenarios

Tracking the Ball

Type: Single Object Tracking (SOT).
Recommended Algorithm: ByteTrack.
Reasoning:

- ByteTrack's ability to utilize low-confidence detections ensures robust tracking of fast-moving objects like the ball.
- Temporal modules, such as TSM, enhance consistency during occlusions or rapid movements.

Tracking Players

Type: Multiple Object Tracking (MOT).
Recommended Algorithms: FairMOT, DeepSORT, MixSort.
Reasoning:

- Players often interact closely, requiring robust re-identification to maintain consistent identities.
- FairMOT integrates detection and re-identification tasks, excelling in crowded environments.
- MixSort leverages appearance and motion cues for improved association in dense scenarios.

Tracking Referees

Type: Multiple Object Tracking (MOT).
Recommended Algorithm: ByteTrack.
Reasoning:

- Referees are typically less dynamic than players but still require reliable tracking.
- ByteTrack's handling of low-confidence detections ensures consistent tracking.

Summary

- **Ball Tracking:** ByteTrack and temporal models like TSM excel in handling the ball's rapid motion and occlusions.
- **Player Tracking:** FairMOT and MixSort provide robust solutions for tracking players in crowded and dynamic sports scenarios.
- **Referee Tracking:** ByteTrack offers reliable tracking for referees, ensuring consistency across frames.

3 Comprehensive Explanation of Tracking Metrics

Object tracking metrics serve to evaluate different aspects of a tracking system, including detection accuracy, identity consistency, and overall performance. Below is a detailed explanation of **MOTA**, **IDF1**, **Recall**, and **HOTA**, highlighting their strengths, limitations, and how they complement each other.

1. MOTA (Multiple Object Tracking Accuracy)

Definition: MOTA evaluates overall tracking performance by combining false negatives (missed detections), false positives (incorrect detections), and identity switches into a single metric:

$$MOTA = 1 - \frac{\sum_t (FN_t + FP_t + IDS_t)}{\sum_t GT_t}$$

Strengths:

- **Holistic:** Incorporates detection and identity errors into one score.
- **Simplicity:** Provides a straightforward, single-number evaluation.

Limitations:

- **Bias Toward Detection:** Penalizes false negatives more heavily than identity switches, making it less reliable for evaluating identity consistency.
- **No Identity Tracking Details:** It doesn't focus on how well object identities are maintained.
- **Over-Simplification:** Combines different error types into one score, potentially masking specific weaknesses.

How Other Metrics Address MOTA's Limitations:

- **IDF1:** Directly focuses on identity consistency, complementing MOTA's lack of identity tracking evaluation.
- **HOTA:** Balances detection and association accuracy, providing a more nuanced evaluation than MOTA.

2. IDF1 (ID F1 Score)

Definition: IDF1 focuses on identity consistency by evaluating how well object identities are preserved throughout the video. It is the harmonic mean of Identification Precision (IDP) and Identification Recall (IDR):

$$IDF1 = \frac{2 \cdot IDTP}{2 \cdot IDTP + IDFP + IDFN}$$

Strengths:

- **Focus on Identity:** Captures identity preservation, a critical aspect in object tracking.
- **Balances Precision and Recall:** Ensures a fair evaluation of both over-detection and missed identities.

- **Robust to Crowded Scenes:** Handles identity tracking better in complex scenarios.

Limitations:

- **Ignores Detection Quality:** Does not directly account for false positives or false negatives in detection.
- **Simplifies Complex Scenarios:** In crowded environments, identity switches may not be fully captured.
- **IoU Dependency:** Uses a fixed IoU threshold, which might not suit all object sizes or scenarios.

How Other Metrics Address IDF1's Limitations:

- **MOTA:** Evaluates detection errors (false positives, false negatives), complementing IDF1's identity-only focus.
- **HOTA:** Incorporates both detection and association, addressing IDF1's lack of detection evaluation.

3. Recall

Definition: Recall measures the percentage of ground truth objects successfully detected and tracked:

$$Recall = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

Strengths:

- **Completeness:** Measures how well all objects are detected.
- **Simple and Intuitive:** Easy to interpret and compute.

Limitations:

- **No Penalization for False Positives:** A tracker with many incorrect detections can still achieve high recall.
- **Focus on Detection Only:** Does not evaluate identity consistency or tracking performance.

How Other Metrics Address Recall's Limitations:

- **MOTA:** Incorporates both false positives and identity switches to evaluate overall performance.
- **IDF1:** Adds identity evaluation, ensuring that detected objects are consistently tracked.

4. HOTA (Higher Order Tracking Accuracy)

Definition: HOTA is a modern metric that balances detection accuracy ($DetA$) and association accuracy ($AssA$) to provide a comprehensive evaluation:

$$HOTA_\alpha = \sqrt{\text{DetA}_\alpha \cdot \text{AssA}_\alpha}$$

Strengths:

- **Balanced Evaluation:** Combines detection and identity association into a single score.
- **Multiple IoU Thresholds:** Evaluates performance across a range of IoU values, making it more adaptable.
- **Handles Long-Term Tracking:** Evaluates global associations over the entire sequence.

Limitations:

- **Not Ideal for Online Tracking:** Requires future data, making it unsuitable for real-time applications.
- **No Penalty for Fragmentation:** Does not consider how smoothly trajectories are tracked.

How Other Metrics Address HOTA's Limitations:

- **MOTA:** Penalizes identity switches and provides a simple evaluation for online tracking scenarios.
- **IDF1:** Captures detailed identity preservation, which complements HOTA's focus on balanced evaluation.

How These Metrics Complement Each Other

- **MOTA + IDF1:** MOTA evaluates overall tracking accuracy, but IDF1 adds detailed identity consistency evaluation.
- **MOTA + Recall:** MOTA captures both detection and identity errors, while Recall ensures completeness of detection.
- **HOTA + MOTA/IDF1:** HOTA balances detection and identity consistency, addressing the limitations of MOTA and IDF1 by providing a holistic view.
- **IDF1 + Recall:** IDF1 evaluates identity tracking, while Recall ensures that all objects are detected, creating a complementary evaluation.

Conclusion: Each metric has unique strengths and limitations. A robust evaluation framework should include multiple metrics (e.g., MOTA, IDF1, Recall, and HOTA) to ensure a comprehensive analysis of tracking systems, balancing detection, identity consistency, and global tracking performance.

4 Review of Tracking-Specific Datasets for Sports Scenarios

1. Overview of SportsMOT Dataset

The SportsMOT dataset is specifically designed for multi-object tracking (MOT) in sports, making it highly relevant for projects focusing on player tracking and analysis. Below are its key features:

- **Purpose:** To enable evaluation of multi-object tracking algorithms in challenging sports scenarios.
- **Core Features:**
  - **Diverse Sports Scenarios:** SportsMOT includes high-paced team sports like soccer and basketball, with frequent player interactions and occlusions.
  - **Detailed Annotations:** Provides frame-by-frame bounding boxes, player identities, and team affiliations, making it ideal for tasks requiring identity preservation.
  - **Challenging Environments:**
    - Dense player overlaps and occlusions.
    - Rapid changes in player movement and direction.
    - Varied camera angles and perspectives.

2. Relevance of SportsMOT for The Project

Sports scenarios present unique challenges for tracking algorithms, including:

- **High-Speed Movement:** Players frequently accelerate, decelerate, or change direction, requiring advanced motion modeling.
- **Frequent Occlusions:** Players and objects (e.g., balls, referees) often block each other, complicating identity tracking.
- **Identity Preservation:** Maintaining consistent player identities over time is critical for deeper analysis of tactics, performance, and interactions.

3. Exploring Complementary Datasets

While SportsMOT is highly relevant, combining it with other datasets can enrich your project and provide varied challenges for evaluation. Here are some additional options:

1. **SoccerNet:**
   - Focuses on soccer matches with player tracking and action detection.
   - Provides bounding boxes, team annotations, and event labels.
   - Link: https://www.soccer-net.org/
2. **PoseTrack:**
   - A dataset for human pose estimation and multi-person tracking.
   - Useful for analyzing player positions and movements in sports.

- Link: https://posetrack.net/

3. **NBA-Ish Dataset:**
   - Designed for basketball tracking, capturing fast-paced movements and tight interactions between players.
   - Includes ball tracking and detailed play analysis.

4. **KITTI MOT Dataset (for general MOT testing):**
   - While not sports-specific, KITTI MOT provides challenging multi-object tracking scenarios in real-world outdoor scenes.

5 Challenges Related to Tracking and Their Impacts

Object tracking systems face real-world challenges such as occlusion, scale variation, and illumination changes that can significantly affect their performance and evaluation metrics. Below is a comprehensive explanation of these challenges, their impacts, and possible solutions.

1. Occlusion

Definition: Occlusion occurs when one object blocks another partially or fully, making it difficult for the tracker to detect or correctly identify the occluded object.

Impact on Tracking:

- **Metrics Affected:**
  - **MOTA:** Increased false negatives (FN) as occluded objects may not be detected, reducing overall accuracy.
  - **IDF1:** Higher identity switches (IDSW) due to occluded objects being assigned new identities when they reappear.
  - **HOTA:** Both detection accuracy ($DetA$) and association accuracy ($AssA$) degrade due to missed detections and broken identity assignments.
- **Real-World Example:** Players frequently overlap during crowded scenarios in sports, such as a corner kick in soccer, leading to temporary occlusions.

Solutions:

- Use **Re-Identification (Re-ID) modules** to restore identities of occluded objects.
- Incorporate **temporal tracking models** to predict object trajectories through occluded regions.
- Leverage **3D tracking techniques** that account for depth to reduce the impact of overlaps.

2. Scale Variation

Definition: Scale variation occurs when objects change size significantly due to camera zoom, distance, or perspective. For example, players appear smaller when further from the camera and larger when closer.

Impact on Tracking:

- **Metrics Affected:**
  - **Recall:** Distant, smaller objects might be missed, leading to increased false negatives and lower recall.
  - **MOTA:** Increased false positives (FP) and false negatives (FN) as detection errors occur at varying scales.
  - **HOTA:** Detection accuracy ($DetA$) decreases as smaller objects are harder to detect reliably.
- **Real-World Example:** In soccer, players on the far end of the field may not be tracked as accurately as players near the camera.

Solutions:

- Use **multi-scale feature extraction** in detection models to handle varying object sizes.
- Train models on **datasets with diverse scales** to improve generalization.
- Implement **adaptive feature pyramids** that dynamically focus on objects of different scales.

3. Illumination Change

Definition: Illumination changes, such as shadows, glare, or varying lighting conditions, can reduce the visibility of objects and disrupt detection and tracking performance.

Impact on Tracking:

- **Metrics Affected:**
  - **MOTA:** Increased FP and FN due to inconsistent detection under poor lighting.
  - **IDF1:** Identity consistency declines as objects in shadow or glare may be misidentified or missed altogether.
  - **HOTA:** Both $DetA$ and $AssA$ degrade due to feature inconsistencies caused by lighting variations.
- **Real-World Example:** In outdoor sports, sudden shadows from moving clouds or glare from sunlight can obscure players or objects, making tracking harder.

Solutions:

- Use **data augmentation** to simulate brightness, contrast, and shadow variations during training.
- Employ **illumination-invariant feature extractors**, such as gradient-based or edge-based descriptors.
- Incorporate **adaptive normalization techniques** to reduce the effect of lighting changes.

6 Data Loader Class Implementation

I uploaded the notebook, next to the PDF file.

7 Challenges in Tracking Algorithms and Recommendations for Improvement

Object tracking in sports presents unique challenges due to dynamic environments, frequent occlusions, and rapid object movements. Below is a comprehensive analysis of these challenges and strategies to enhance tracking performance.

7.A Challenges in Object Tracking for Sports

Occlusion and Identity Loss

Problem: Players and objects (e.g., the ball) frequently overlap, causing identity switches, tracking drift, and re-identification errors.
Impact: Reduces tracking accuracy, especially in crowded areas like goalposts and free kicks.

Scale Variation

Problem: Objects appear at different scales due to zoom, perspective changes, and varying player-camera distances.
Impact: Leads to bounding box mismatches and feature degradation.

Illumination and Environmental Variations

Problem: Shadows, glare, and varying lighting conditions alter object appearance, making detections unstable.
Impact: Increases false positives and false negatives, affecting detection confidence.

Dynamic Backgrounds and Scene Complexity

Problem: Background elements (crowd, advertisements, moving referees) interfere with object tracking.
Impact: Incorrect object associations and false detections.

High-Speed Motion and Motion Blur

Problem: Players and balls move at high velocities, often exceeding 30 km/h, causing motion blur.
Impact: Leads to tracking failures, especially for fast-moving objects like the ball.

Lack of Robust Evaluation Metrics

Problem: Metrics like MOTA and IDF1 fail to fully capture tracking performance in complex sports environments.
Impact: Hard to assess real-world tracking accuracy comprehensively.

7.B Recommendations for Performance Improvement

Occlusion Handling and Re-Identification

- **Occlusion-Aware Models:** Use trajectory prediction and temporal modeling to restore lost object identities.

- **Re-ID Enhancements:** Incorporate appearance-based deep learning models (e.g., transformer-based Re-ID) to track occluded players.
- **Graph-Based Tracking:** Model spatial relationships between players to improve identity re-association.

Scale Adaptation and Multi-Scale Feature Learning

- **Feature Pyramid Networks (FPNs):** Improve detection of small and distant objects.
- **Adaptive Bounding Boxes:** Adjust box sizes dynamically based on object motion and scale.

Illumination-Invariant Tracking

- **Lighting-Augmented Training:** Train models with brightness, contrast, and shadow variations.
- **Normalization Techniques:** Apply histogram equalization and adaptive contrast enhancement.

Scene Awareness and Background Filtering

- **Dynamic Background Subtraction:** Use context-aware filtering to separate players from moving crowds.
- **Object-Specific Embeddings:** Assign team colors and uniform-based ID features for better player tracking.

High-Speed Motion Prediction

- **Kalman Filters + LSTMs:** Predict ball/player trajectories across frames.
- **Higher Frame Rate Processing:** Train models on 50-60 FPS data for smoother tracking.

Robust Evaluation Metrics

- **HOTA (Higher Order Tracking Accuracy):** Balances detection, association, and localization quality.
- **Hybrid Evaluation:** Combine traditional (MOTA, IDF1) and qualitative analysis for more comprehensive assessment.

Temporal Modeling for Long-Term Tracking

- **Transformer-Based Tracking:** Leverage self-attention mechanisms to model long-term player interactions.
- **RNN-Based Motion Learning:** Use GRUs/LSTMs to improve trajectory forecasting.

Synthetic Data Augmentation

- **Simulation-Based Training:** Generate synthetic sports data to improve generalization in rare scenarios.
- **Domain Adaptation Techniques:** Fine-tune models with real and synthetic datasets to improve robustness.

Conclusion

By addressing these challenges with adaptive tracking models, multi-scale feature extraction, advanced motion prediction, and robust evaluation metrics, sports tracking algorithms can be significantly improved. These enhancements will reduce identity switches, improve tracking accuracy, and provide deeper insights into player performance and game dynamics.