Multiple Object Tracking for First Robotics Competition

Gabriel (94936730)

Felix (94936730)

Kevin (94936730)

Josh ((94936730)

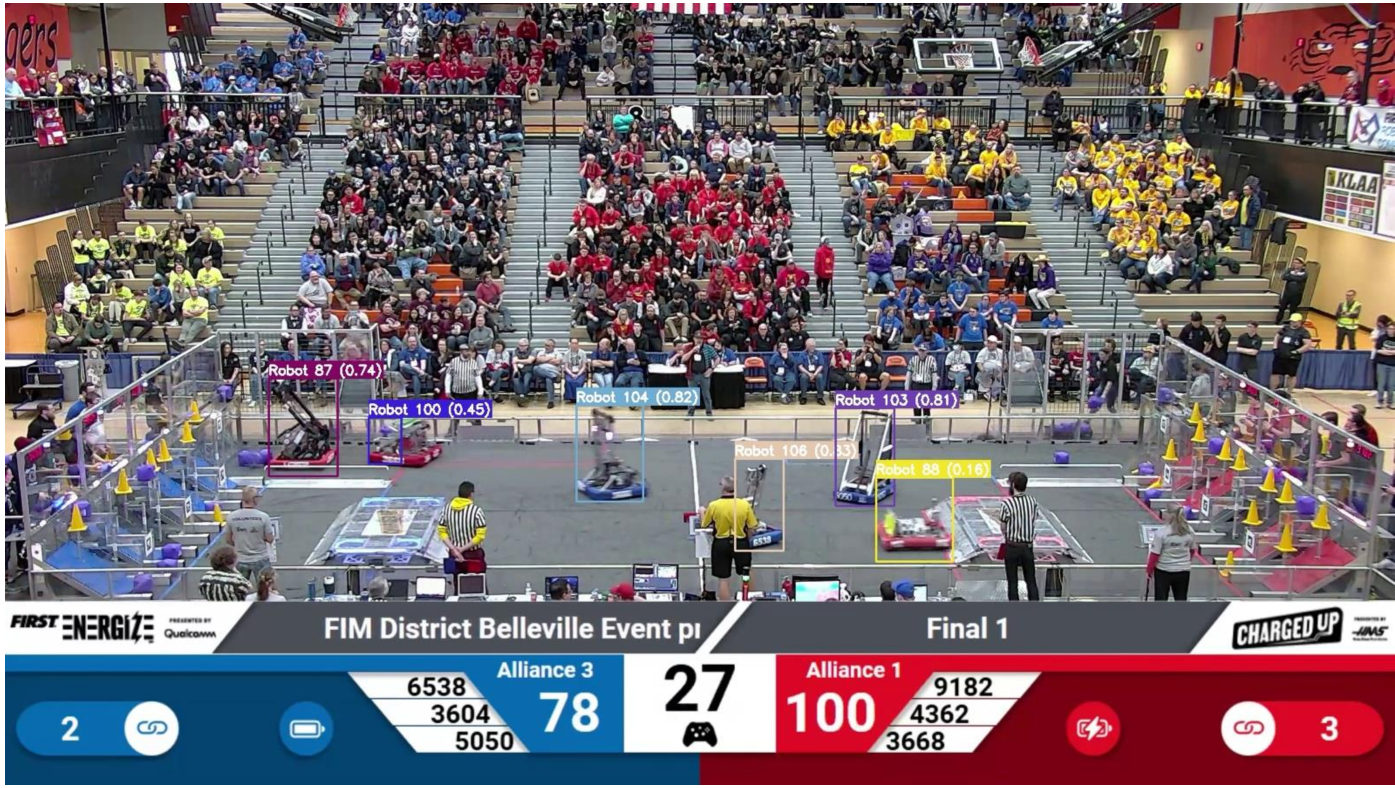
Course: EECS 442 (Fall 2025)

FRC®

FIRST® Robotics Competition

Introduction

- **First Robotics** is a competition where teams build robots to compete in a game that is new every year
- **Problem:** Currently Scoring of Roboter Performance is done manually
- **Solution:** Implement MOT Solution to track Roboter automatically
- **Research Objective:**
  - Fine Tuning of Yolo Model on FRC data
  - Implementation and Evaluation of SOTA MOT models on FRC Videos



Regular MOT[1]

- Serves as the baseline MOT model for comparison
- Extension of the classic SORT tracking algorithm
  - adds a **deep learning appearance model for tracking**
  - Uses a Kalman filter for motion prediction (based on velocity)
  - Applies the Hungarian algorithm for detection–track assignment
- Visual re-identification (Re-ID): **Extracts appearance features using a pre-trained ResNet CNN**
- **Maintains a feature gallery** for each tracked robot which enables robust re-identification after occlusion

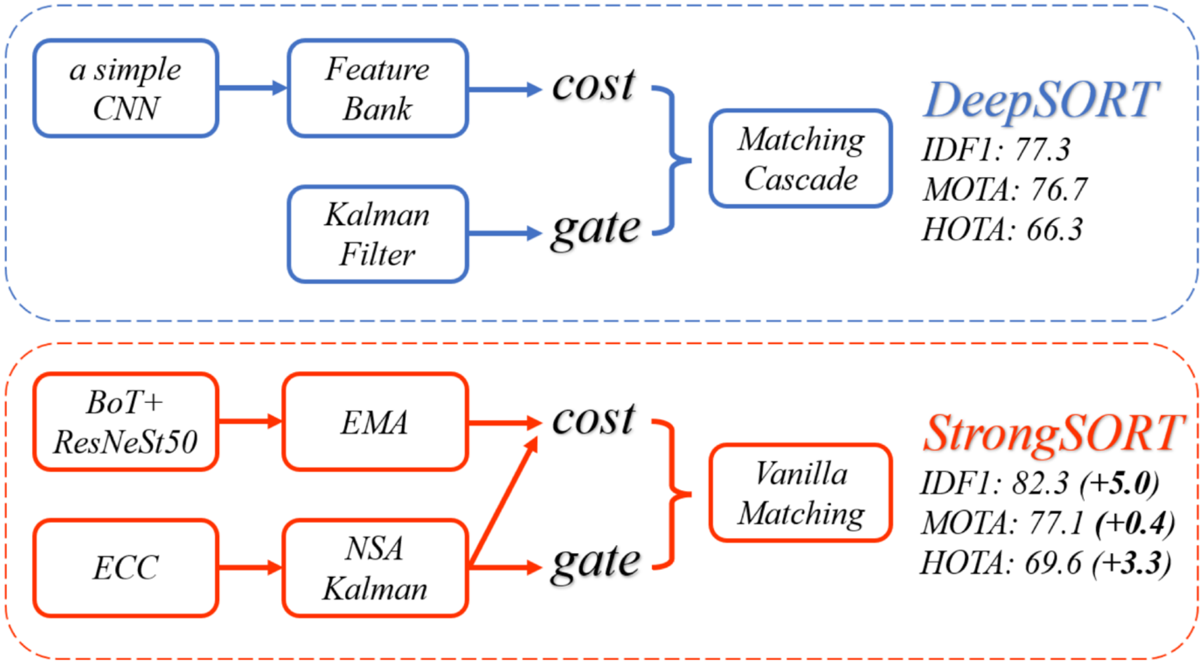
→ Maintain correct tracks of fast-moving robots

→ Reduce false new tracks after temporary disappearances

1 Cf. Buric, Matija, Marina Ivasic-Kos, and Miran Pobar. "Player tracking in sports videos." *2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*. IEEE, 2019.

StrongSort[2]

- Improved Appearance Feature Extraction
  - Standard MOT uses a shallow CNN for feature extraction
  - **StrongSORT uses Bag of Tricks (BoT) with ResNeSt50**
  - Provides more discriminative appearance features
  - Especially important in our use case:
    - Many robots share similar standard designs
    - Accurate re-identification between similar robots is critical
- Noise Scale Adaptive Kalman Filter
  - Uses a **confidence-aware Kalman Filter**
  - Detection confidence influences filter weighting:
    - Low confidence detection → trust motion prediction more
    - High confidence detection → trust observation more
- Improves robustness to Noisy detections and Temporary tracking uncertainty



2 Cf. Du, Yunhao, et al. "Strongsort: Make deepsort great again." *IEEE Transactions on Multimedia* 25 (2023): 8725-8737.

OC-SORT[3]

- Yolo Model for Detections / Extension of the classic SORT tracking algorithm
- **ORU: Observation Re-Update Mechanism**
  - Activated when a track is re-associated after being untrack
  - Creates a virtual trajectory between:
    - Last visible position (before occlusion)
    - First re-detected position (after occlusion)
  - Generates virtual observations for each occluded frame -> re-update the Kalman Filter state

$$
\textit{re-update} \begin{cases} \mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^\top(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^\top + \mathbf{R}_t)^{-1} \\ \hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\tilde{\mathbf{z}}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}) \\ \mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1} \end{cases}
$$

- **Observation-Centric Momentum (OCM)**
  - Uses Intersection over Union (IoU) for standard data association
  - Adds a momentum-based motion consistency factor:
    - Angle of Track: direction from last two observations
    - Angle of Intention: direction from last observation to new detection

$$
C(\hat{\mathbf{X}}, \mathbf{Z}) = C_{\mathrm{IoU}}(\hat{\mathbf{X}}, \mathbf{Z}) + \lambda C_v(\mathcal{Z}, \mathbf{Z}),
$$

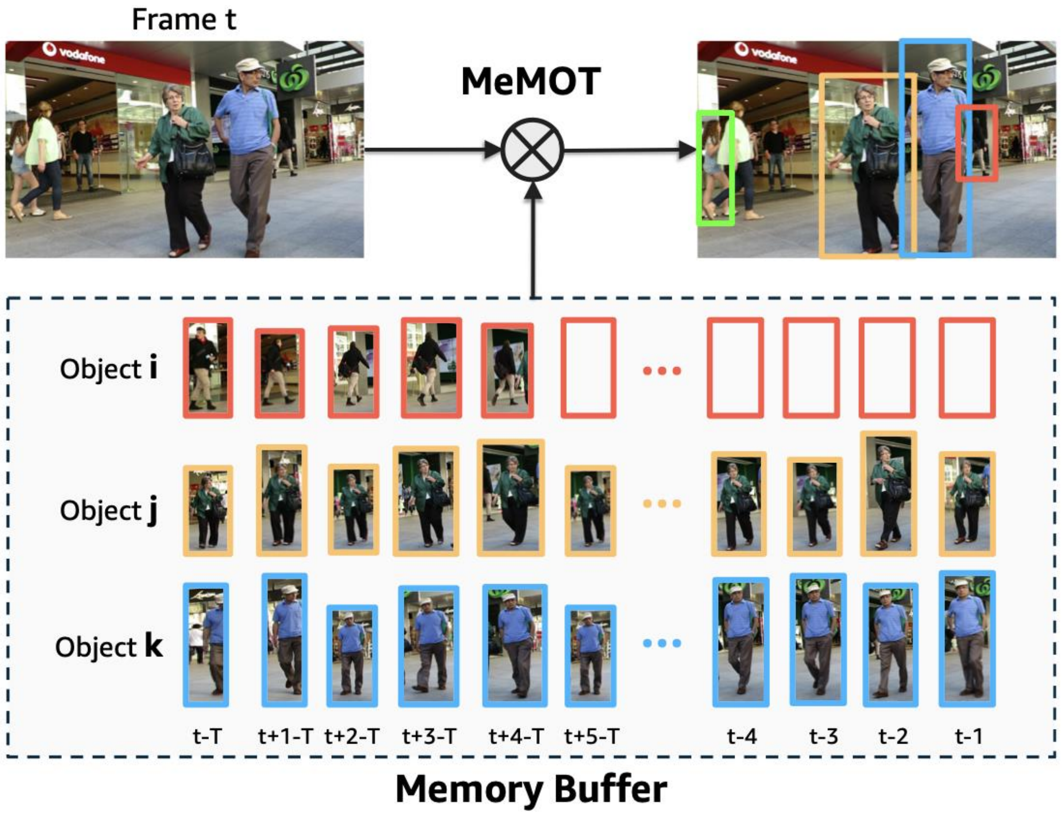
$$
\Delta\theta = |\theta^{\text{track}} - \theta^{\text{intention}}|
$$

- Improves tracking stability and accuracy after occlusions

3 Cf. Cao, Jinkun, et al. "Observation-centric sort: Rethinking sort for robust multi-object tracking." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023.

MeMOT[4]

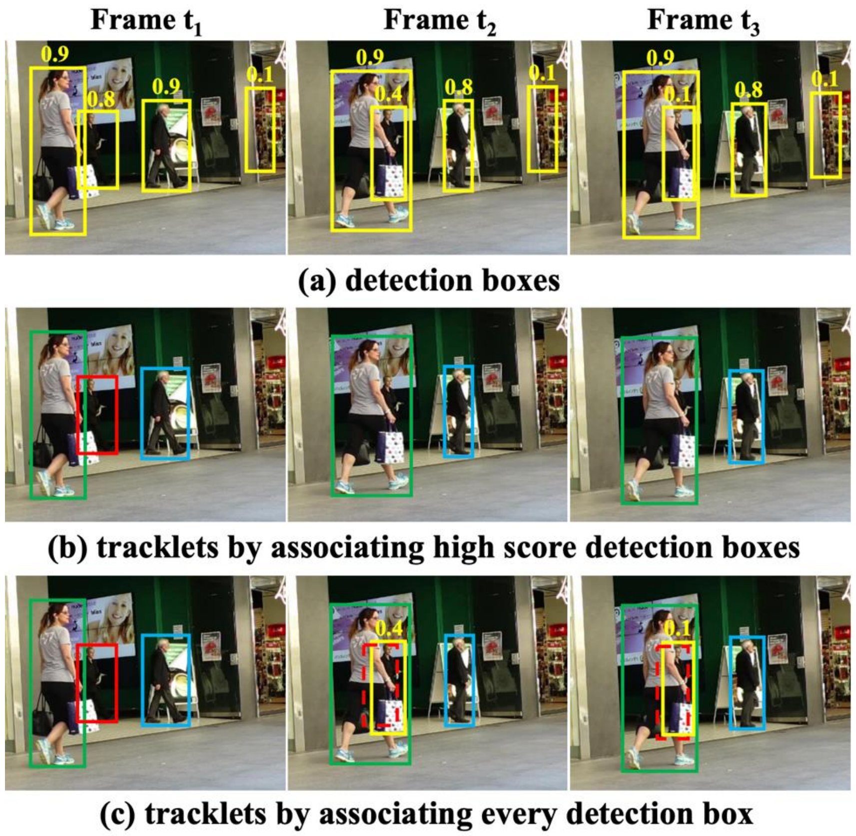
- Uses YOLO for detections
- Adds **MeMOT-style memory** with appearance embeddings
- Maintains stable robot identities across frames
- Reduces ID switches compared to simple MOT methods
- Strengths:
  - Runs in real time
  - Handles minor occlusions and jitter well
  - Provides good identity consistency when detections are clean
- Limitations
  - Strong dependence on YOLO detection accuracy
  - Long occlusions can break tracks
  - Identity flips when robots overlap closely
  - Less robust than full MeMOT (no transformer modules)



4 Cf. Cai, Jiarui, et al. "Memot: Multi-object tracking with memory." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022..

ByteTrack[5]

- Algorithm built on YOLO detections
- Uses a **two-stage association** strategy:
  - Matches high-confidence detections first
  - Uses low-confidence detections to recover missed tracks
- Applies a Kalman filter for motion prediction (constant velocity)
- Uses the Hungarian algorithm for optimal track–detection assignment
- Strengths:
  - Runs in real time and is computationally efficient
  - Handles occlusions well without relying on appearance features
- Well-suited for FRC footage:
  - Lower video quality
  - Fast-moving robots
  - Frequent occlusions

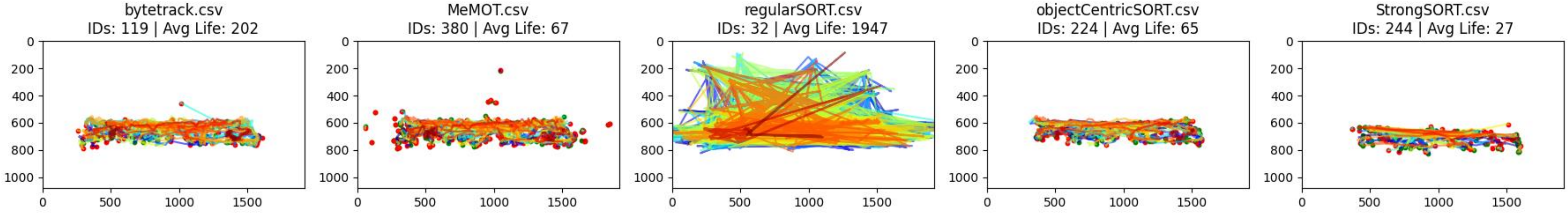


5 Cf. Zhang, Yifu, et al. "Bytetrack: Multi-object tracking by associating every detection box." *European conference on computer vision*. Cham: Springer Nature Switzerland, 2022

Model Evaluation

Tracking Performance Comparison

Model	Total IDs	Avg Life (Frames)	Max Life	Gap Count
RegularSORT	32	1947	3546	30
ByteTrack	119	202	1088	86
MeMOT	380	67	2041	123
ObjectCentricSORT	224	65	748	149
StrongSORT	244	27	351	174



Literature Review & Sources

Buric, Matija, Marina Ivasic-Kos, and Miran Pobar. "Player tracking in sports videos." *2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*. IEEE, 2019.

Cao, Jinkun, et al. "Observation-centric sort: Rethinking sort for robust multi-object tracking." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023.

Cai, Jiarui, et al. "Memot: Multi-object tracking with memory." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022.

Du, Yunhao, et al. "Strongsort: Make deepsort great again." *IEEE Transactions on Multimedia* 25 (2023): 8725-8737.

Zhang, Yifu, et al. "Bytetrack: Multi-object tracking by associating every detection box." *European conference on computer vision*. Cham: Springer Nature Switzerland, 2022

First Robotics Competition: https://www.firstinspires.org/programs/frc/