Ball Trajectory Inference from Multi-Agent Sports Contexts Using Set Transformer and Hierarchical Bi-LSTM

The 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2023)

Hyunsung Kim
Fitogether Inc.
Seoul, South Korea

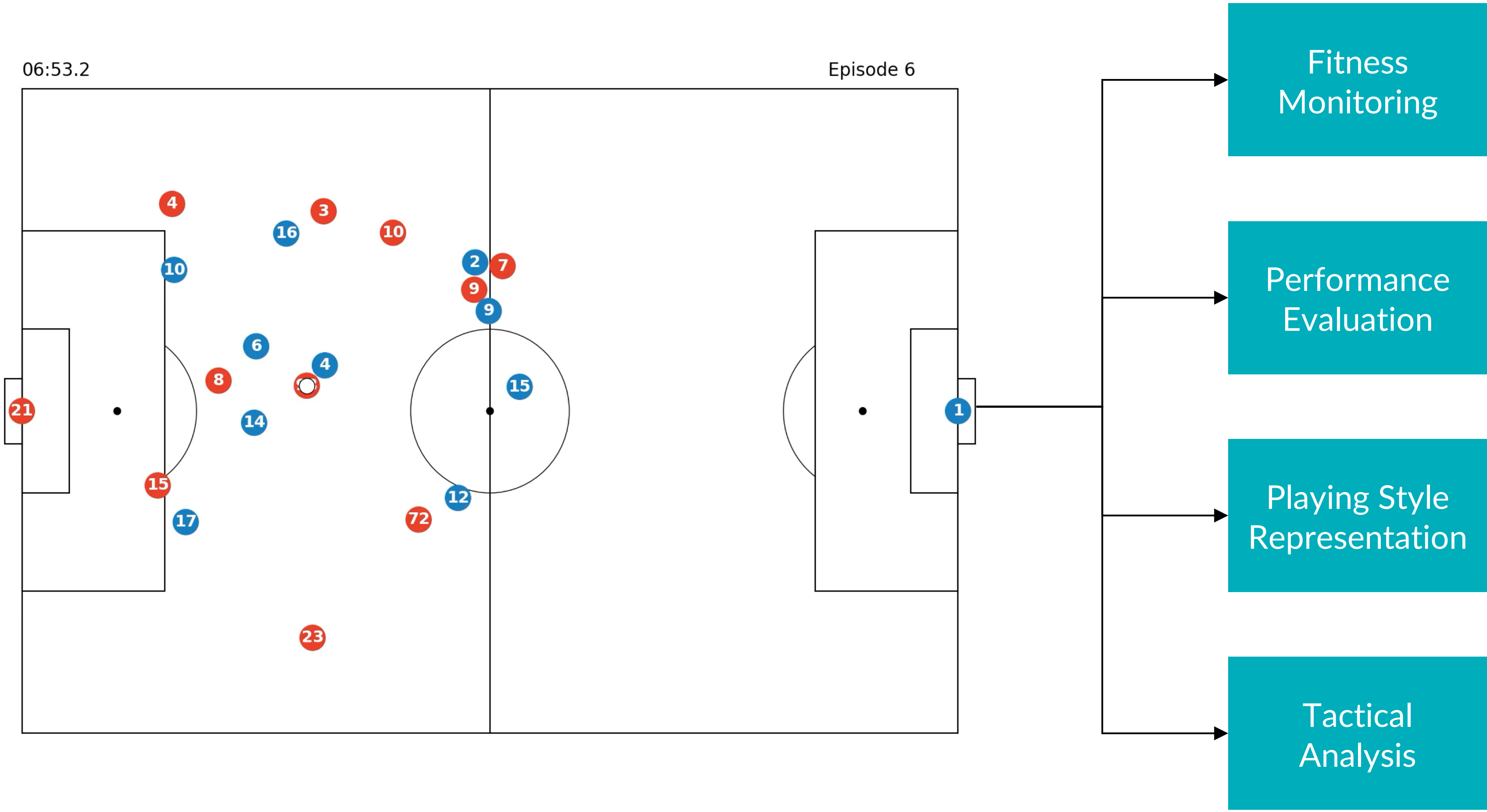
Han-Jun Choi
Kangwon National University
Chuncheon, South Korea

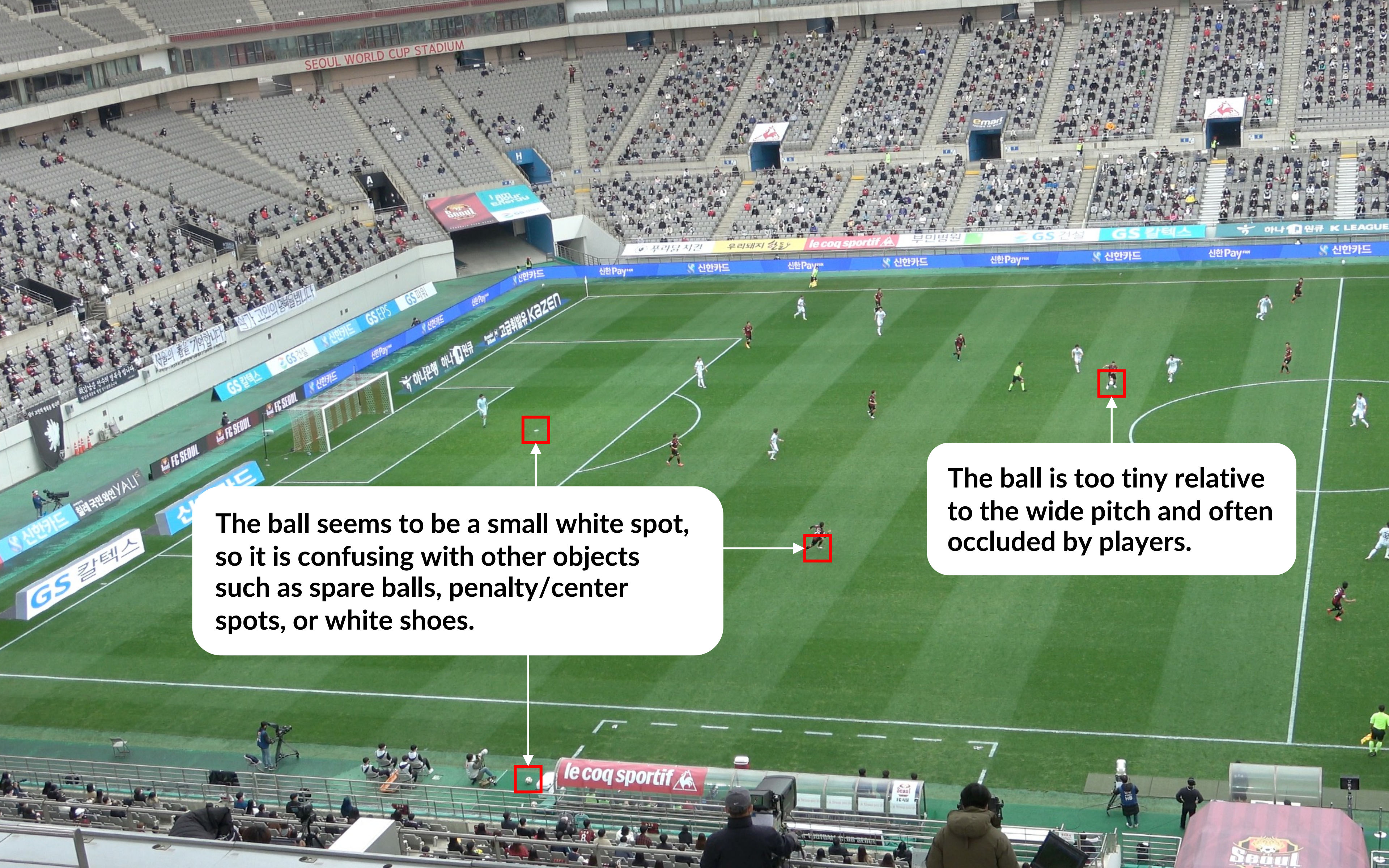
Chang Jo Kim
Fitogether Inc.
Seoul, South Korea

Jinsung Yoon
Fitogether Inc.
Seoul, South Korea

Sang-Ki Ko
Kangwon National University
Chuncheon, South Korea

Introduction: Player and Ball Tracking Data in Team Sports




SEOUL WORLD CUP STADIUM
The ball seems to be a small white spot, so it is confusing with other objects such as spare balls, penalty/center spots, or white shoes.
The ball is too tiny relative to the wide pitch and often occluded by players.
le coq sportif
FC SEOUL
GS 칼텍스
신한카드
kazen

Framework for Context-Aware Ball Trajectory Inference

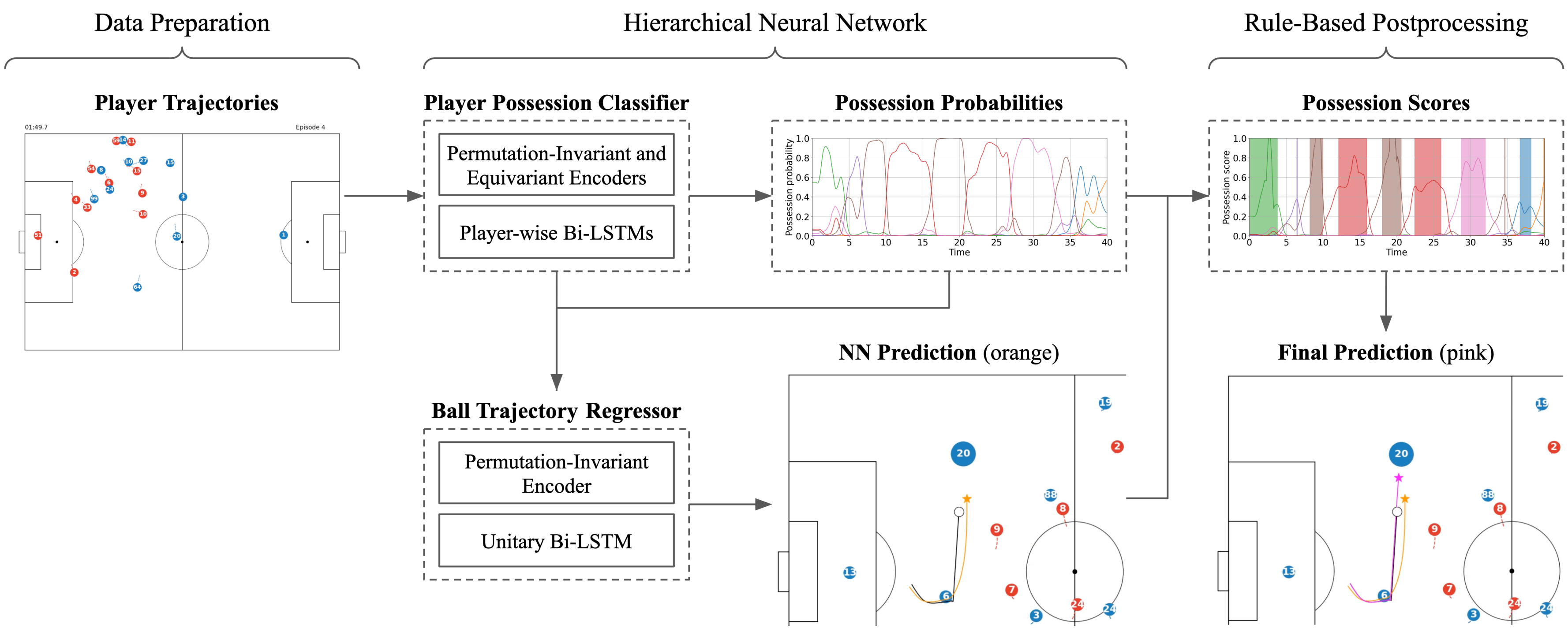
Predicting the ball trajectory from player trajectories as a cost-efficient alternative to video-based ball tracking

Player possession classification (PPC)

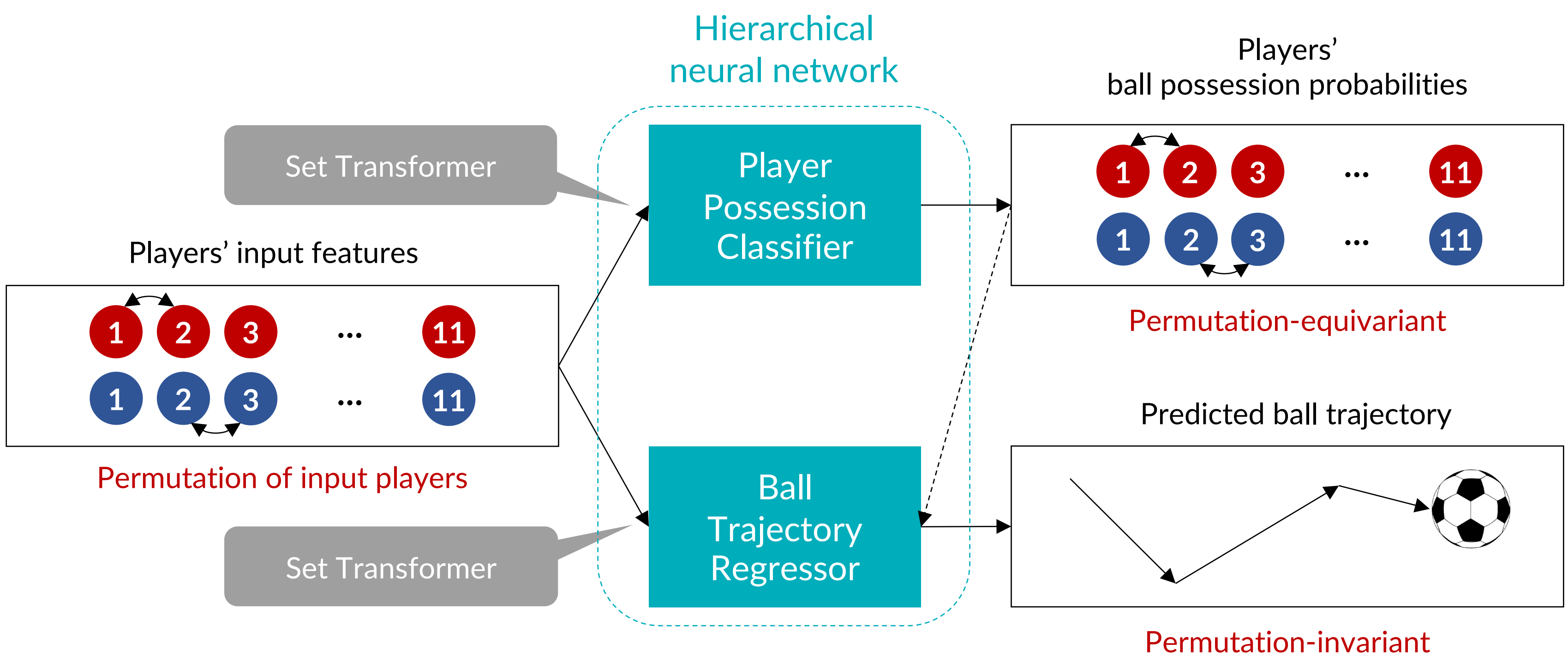
Given a set $X^P_{1:T} = \{\mathbf{x}^p_{1:T}\}_{p \in P}$ of player trajectories, find the player q_t that possesses the ball at each time t.

Ball trajectory regression (BTR)

Given a set $X^P_{1:T} = \{\mathbf{x}^p_{1:T}\}_{p \in P}$ of player trajectories, estimate the ball trajectory $\mathbf{y}_{1:T}$.

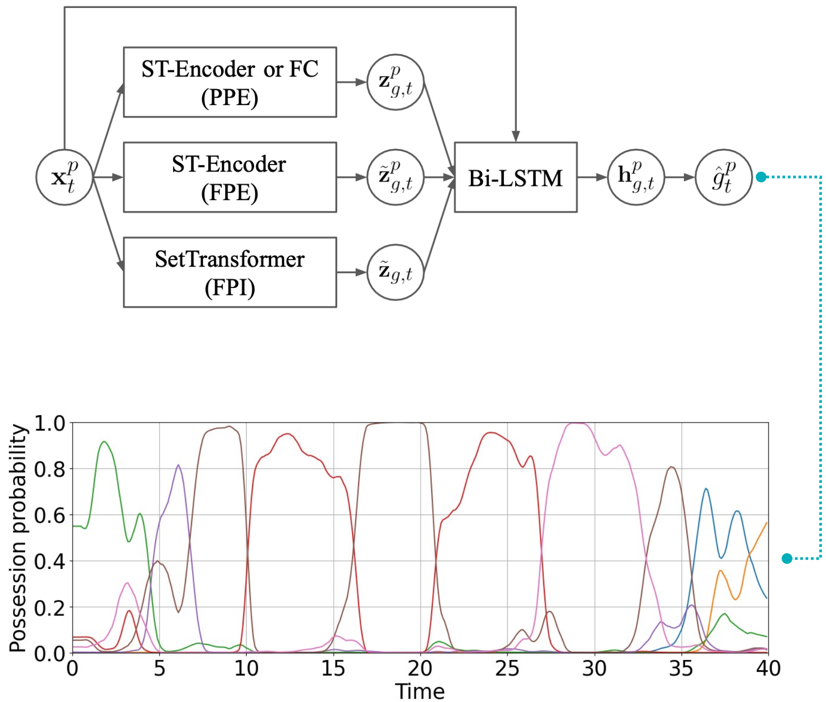


Permutation-Invariant/Equivariant Context Representations



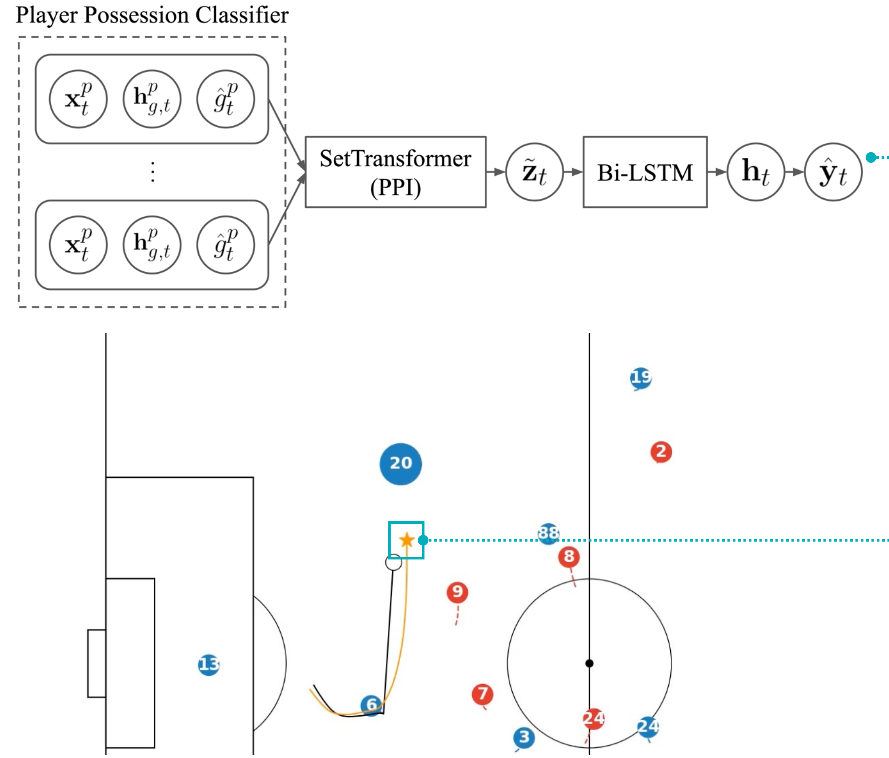
Detailed Model Architectures

Player Possession Classifier



Players' ball possession probabilities

Ball Trajectory Regressor



Predicted ball trajectory

Loss Function

Cross entropy loss for the accuracy of PPC

$$\mathcal{L}^{\mathrm{CE}}(\hat{\mathbf{g}}_{1:T}, \mathbf{g}_{1:T}) = -\frac{1}{T}\sum_{t=1}^{T}\sum_{k=1}^{K} g_t^k \log \hat{g}_t^k$$

MSE loss for the accuracy of BTR

$$\mathcal{L}^{\mathrm{MSE}}(\hat{\mathbf{y}}_{1:T}, \mathbf{y}_{1:T}) = -\frac{1}{T}\sum_{t=1}^{T}\|\hat{\mathbf{y}}_t - \mathbf{y}_t\|_2^2$$

Reality loss for naturalness of ball trajectories

$$\mathcal{L}^{\mathrm{Real}}(\hat{\mathbf{y}}_{1:T}; X_{1:T}^{p}) = -\frac{1}{T-2}\sum_{t=1}^{T-1}\tanh\theta_t \cdot d_t$$

where

$$\theta_t = \arccos\left(\frac{\mathbf{v}_t \cdot \mathbf{v}_{t+1}}{\|\mathbf{v}_t\|\|\mathbf{v}_{t+1}\|}\right)$$

is the course angle of the trajectory at t with the direction $\mathbf{v}_t = \hat{\mathbf{y}}_t - \hat{\mathbf{y}}_{t-1}$ of the ball, and

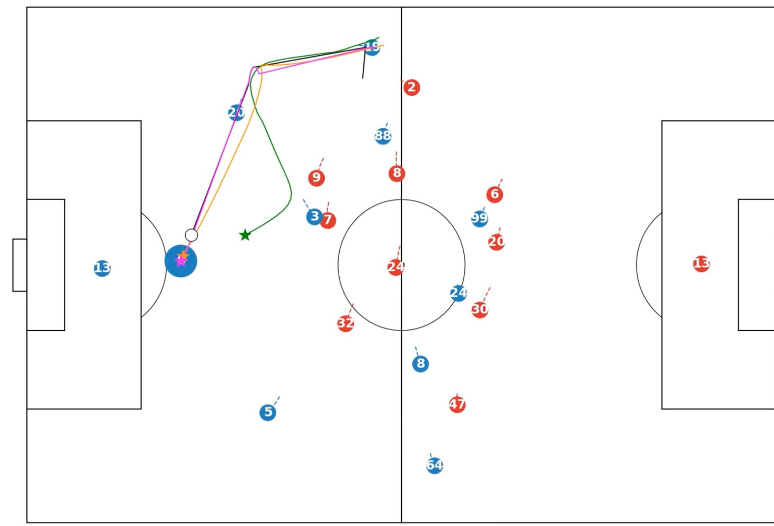
$$d_t = \min_{p \in P}\left\|\hat{\mathbf{y}}_t - \mathbf{x}_t^p\right\|_2$$

is the distance between the ball and the nearest player.

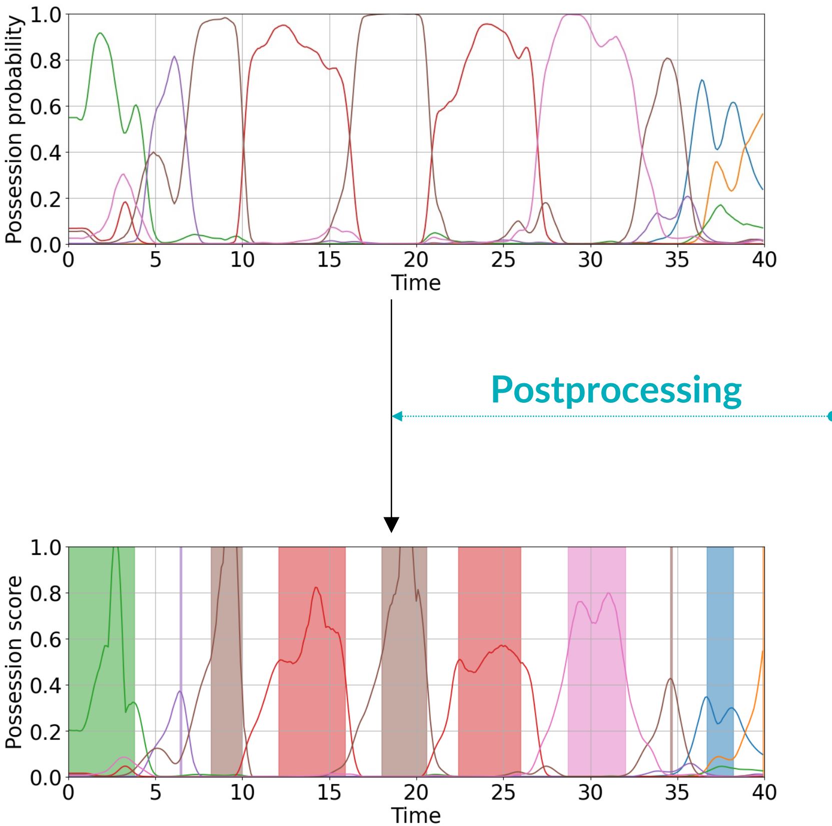
Total loss function

$$\mathcal{L} = \mathcal{L}^{\mathrm{MSE}} + \lambda^{\mathrm{Real}}\mathcal{L}^{\mathrm{Real}} + \lambda^{\mathrm{CE}}\mathcal{L}^{\mathrm{CE}}$$

where the coefficients are empirically set to $\lambda^{\mathrm{Real}} = 1$ and $\lambda^{\mathrm{CE}} = 20$ in our study.



Rule-Based Postprocessing



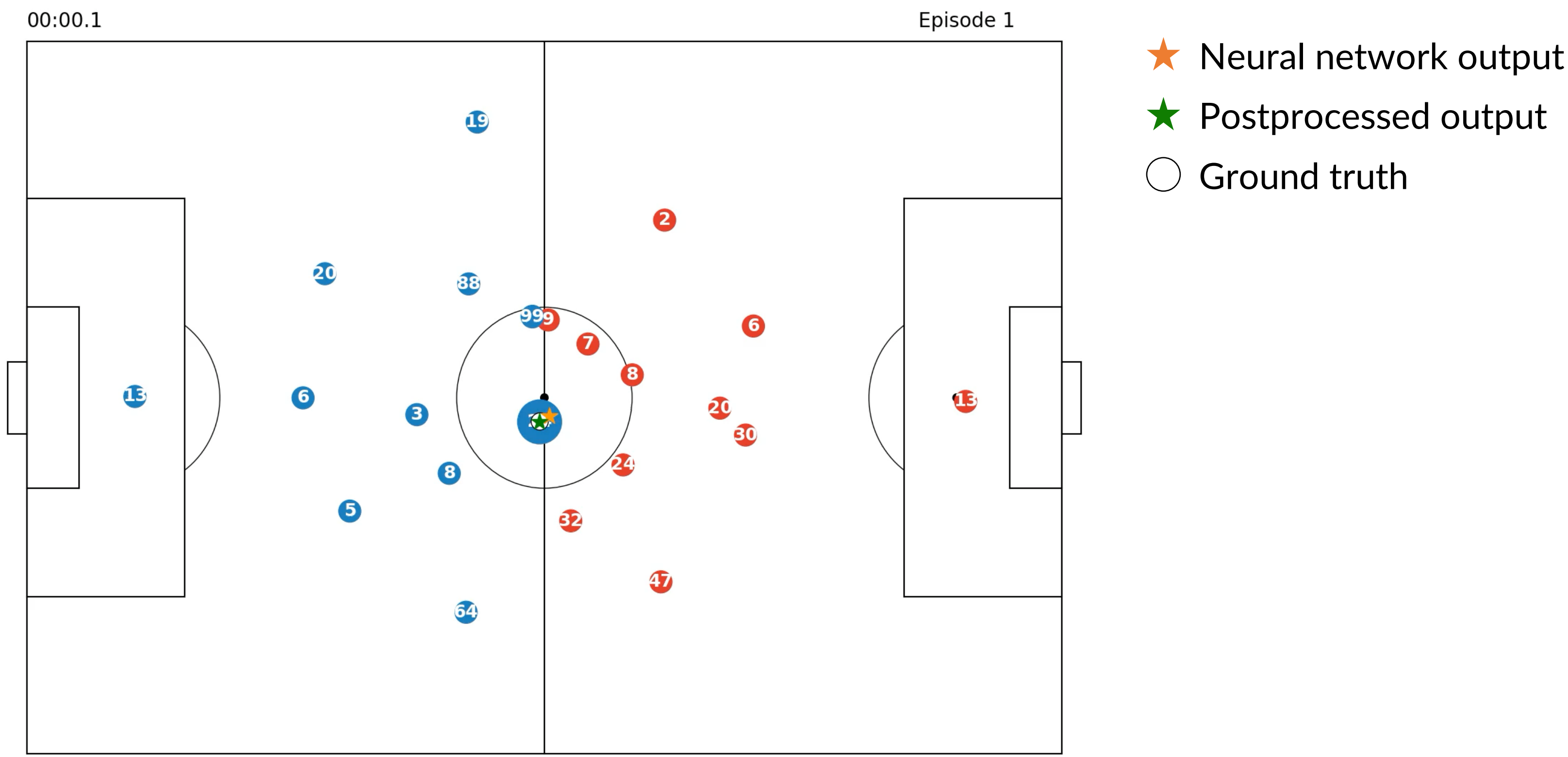
Calculating possession scores

$$\hat{s}_t^p = \frac{\hat{g}_t^p}{\left\| \hat{\mathbf{y}}_t - \mathbf{x}_t^p \right\|_2}$$

Rule-based ball touch detection

- If $\hat{s}_\tau^q = \max_{p \in \tilde{P}} \hat{s}_\tau^p > 0.5$, then q touches the ball at τ.
- If $0.2 < \hat{s}_\tau^q = \max_{p \in \tilde{P}} \hat{s}_\tau^p \leq 0.5$ and $\hat{s}_\tau^q$ is a local maximum of the function $\max_{p \in \tilde{P}} \hat{s}_t^p$ of t, then q touches the ball at τ.
- Otherwise, the ball is moving from one player to another

Visualization of the Resulting Trajectories



Experiments

Dataset

Provider	Split	#. Matches	#. Episodes	#. Frames
Fitogether	Train.	10	533	200,083
	Valid.	2	150	50,708
	Test	3	222	72,191
Metrica	Train.	2	159	71,791
	Valid.	0.5	35	21,120
	Test	0.5	35	21,465

Evaluation metrics

- **Position error (PE):** $-\frac{1}{T}\sum_{t=1}^{T}\|\hat{\mathbf{y}}_t - \mathbf{y}_t\|_2^2$
- **Reality loss (RL):** $-\frac{1}{T-2}\sum_{t=1}^{T-1}\tanh\theta_t \cdot d_t$
- **Player-level possession accuracy (PPA):**
  $$\left|\left\{t: 1 \le t \le T, \arg\max_{p\in\tilde{P}} \hat{g}_t^p = q_t\right\}\right|/T$$
  where q_t denotes the true ball possessor at t.
- **Team-level possession accuracy (TPA):**
  $$\left|\left\{t: 1 \le t \le T, \arg\max_{p\in\tilde{P}} \hat{g}_t^p \in Q_t\right\}\right|/T$$
  where Q_t denotes the team that q_t belongs to.

Experimental Results

- **H-:** A hierarchical model with player-level ball possession as an intermediate target.
- **-RL:** A model trained with reality loss (i.e., $\lambda^{\text{Real}} = 1$).
- **-PP:** A model followed by rule-based postprocessing.

Model	PE	RL	PPA	TPA
VRNN	12.3721	0.7206	-	-
Transformer	7.4727	1.2194	-	-
LSTM	5.4667	0.3392	-	-
H-Transformer	4.9136	0.7987	48.91 %	77.91 %
H-Transformer-RL	4.6933	0.5990	49.52 %	78.01 %
H-Transformer-PP	5.4075	0.0021	52.01 %	78.78 %
H-Transformer-RL-PP	5.4093	0.0021	52.18 %	78.66 %
H-LSTM	3.6886	0.1881	64.01 %	85.26 %
H-LSTM-RL	3.6561	0.1391	64.70 %	85.85 %
H-LSTM-PP	4.1285	0.0016	63.84 %	85.34 %
H-LSTM-RL-PP	4.0719	0.0017	64.32 %	85.34 %

Practical Applications

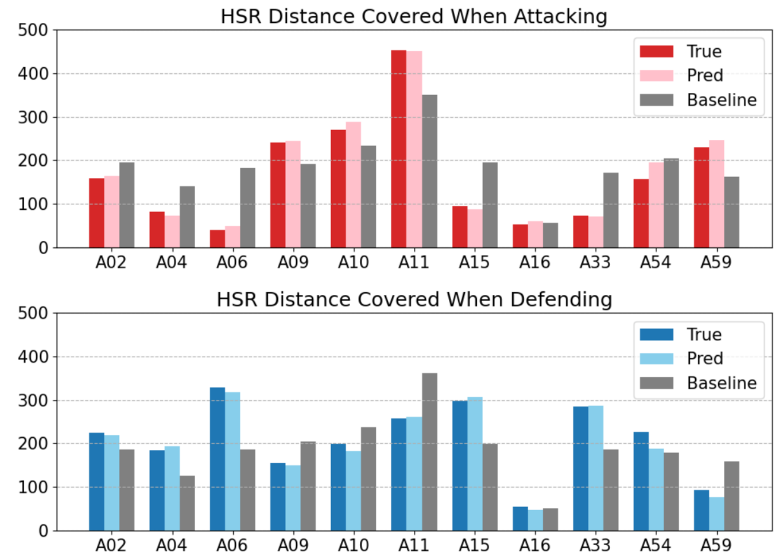
Ball trajectory imputation

Step	Masking	PE	RL	PPA	TPA
Before PP	100 %	5.3536	0.3573	61.91 %	83.96 %
	95 %	3.0018	0.3911	79.51 %	91.54 %
	90 %	2.0939	0.4220	87.11 %	95.04 %
	80 %	1.3052	0.5059	93.13 %	97.41 %
After PP	100 %	5.1440	0.0031	57.29 %	81.06 %
	95 %	3.7990	0.0056	78.71 %	90.73 %
	90 %	2.5913	0.0046	85.92 %	94.50 %
	80 %	1.4137	0.0038	91.25 %	96.67 %

Semi-automated pass annotation

Masking	Pass	Passer	Receiver	#. Passes	#. Receives
100 %	0.3877	0.6039	0.5707	0.7935	0.8150
95 %	0.5972	0.7265	0.7199	0.9031	0.9493
90 %	0.7428	0.8315	0.8208	0.9347	0.9717
80 %	0.8693	0.9149	0.8991	0.9331	0.9746

Possession-wise running performance metrics



Method	Total distance		HSR distance	
	Attacking	Defending	Attacking	Defending
Random	0.0881	0.0771	0.5140	0.2091
Ours	0.0360	0.0340	0.2058	0.1251


00:03.08

Discussion

Contributions

- We propose a framework for inferring the ball trajectory from player trajectories in a sports match as **a cost-efficient alternative to video-based ball tracking**.
- Our framework suggests a way of encoding **permutation-invariant/equivariant nature** of multi-agent sports contexts.
- We introduce a several **practical applications** to be utilized in sports analytics industry.

Future work

- To make the neural network identify whether a player is carrying the ball or it is in transition at a given time in order to **remove the postprocessing step**.
- To construct an end-to-end framework that first obtains reliable fragments of the ball trajectory by **optical tracking** in a match video and then performs **trajectory imputation** leveraging our method.
- To devise other significant and interpretable applications related to match analysis **that utilize the multi-agent attention scores** in Set Transformer.

Ball Trajectory Inference from Multi-Agent Sports Contexts Using Set Transformer and Hierarchical Bi-LSTM

29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2023)

Thank you for listening!

Contact information

Hyunsung Kim (1st author): hyunsung.kim@cml.snu.ac.kr

Sang-Ki Ko (Corr. author): sangkiko@uos.ac.kr