Estimation of Overlapped Tactical Actions from Soccer Match Video

Kento Kuroda[1], Ikuma Uchida[1], Keisuke Fujii[2][a] and Yoshinari Kameda[3][b]
[1]*Graduate School of Science and Technology, University of Tsukuba, Japan*
[2]*Graduate School of Informatics, Nagoya University, Japan*
[3]*Center for Computational Sciences, University of Tsukuba, Japan*

Keywords: Tactical Action, Soccer, Action Estimation.

Abstract: In a soccer match, there may be overlaps of tactical actions performed at a certain point in time, because two teams take different tactical actions and each tactical action has no clear temporal boundary. In this study, we propose a new method for the estimation of overlapped tactical actions from soccer match videos. We enable the estimation of overlapped tactical actions considering exclusive relationships. We achieve this by having the deep learning model learn all tactical actions simultaneously. We validate this method using data from 10 matches. We succeeded in expressing the tactical actions performed at a given time in terms of the strength of several tactical actions.

1 INTRODUCTION

According to Federation Internationale de Football Association (FIFA) coaching manual, soccer is broadly classified into offensive and defensive actions (Barnerat et al., 2000). They are further classified into several tactical actions. In soccer, tactical actions are actions that are performed by the unification of several players to score a goal or defend a goal. People involved in soccer, such as the play-by-play announcers and the coaches, can subjectively estimate tactical actions that are performed in a specific match situation. If we can estimate tactical actions for a specific match situation without having to rely on people involved in soccer, we can treat tactical actions as objective indicators. This would make it possible to represent match situations using the strength of each tactical action performed by the two teams. It is expected that automatic editing of match videos, tactical analysis with more information, and describing match situations will become possible.

Each tactical action can be expressed by the temporal changes of the positional relationships among 22 players and between the ball and the players. There may be overlaps to tactical actions performed at a certain point in time. There are two reasons for this. The first reason is that the two teams will each have their

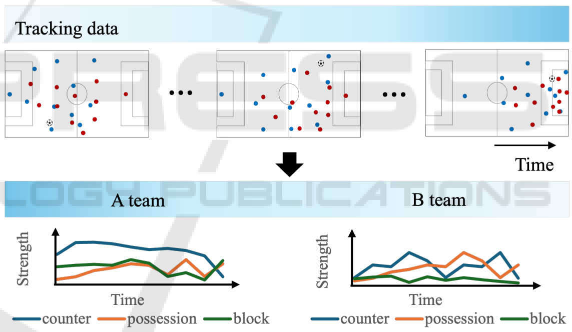

Figure 1: Overview. We estimate tactical actions of the two teams. We use the temporal changes of the positional relationships of the ball and players in tracking data.

tactical actions. The second reason is that the temporal boundaries of tactical actions performed by the same team are unclear. In addition, overlapped tactical actions have exclusive relationships. In soccer, some tactical actions are performed simultaneously, while some are not.

In previous studies, specific tactical actions, such as counter-attacks, were recognized (Sigari et al., 2015; Kobayashi et al., 2012; Fassmeyer et al., 2021; Bauer and Anzer, 2021). In addition, match situations were categorized into five tactical actions (Suzuki et al., 2020). There are two problems with the previous studies. The first problem is that it does not take into account the overlapping of tactical actions. Therefore, the tactical actions that a team performs at a given time are discretely represented by only one

[a] https://orcid.org/0000-0001-5487-4297
[b] https://orcid.org/0000-0001-6776-1267

Kuroda, K., Uchida, I., Fujii, K. and Kameda, Y.
Estimation of Overlapped Tactical Actions from Soccer Match Video.
DOI: 10.5220/0013066300003828
In *Proceedings of the 12th International Conference on Sport Sciences Research and Technology Support (icSPORTS 2024)*, pages 257-264
ISBN: 978-989-758-719-1; ISSN: 2184-3201

in the conventional approaches (Suzuki et al., 2020; Wang et al., 2015). The second problem is that it is difficult to collect a large amount of data when using deep learning to recognize tactical actions.

We propose a new method for the estimation of overlapped tactical actions from soccer match videos. We enable the estimation of overlapped tactical actions considering exclusive relationships. We achieve this by having the deep learning model learn all tactical actions simultaneously. We handcraft a large number of tactical actions based on simple definitions to estimate tactical actions using deep learning. We estimate tactical actions in deep learning by focusing on the temporal changes of the positional relationships among 22 players and between the ball and the players. We estimate eight types of tactical actions, including offensive and defensive tactical actions. The four offensive tactical actions are long counter, short counter, opposition half possession, and own half possession. The four defensive tactical actions are counter-press, high press, middle press, and block. We validated the estimation of overlapped tactical actions considering exclusive relationships by using 10 matches. We succeeded in expressing the tactical actions performed at a given time in terms of the strength of several tactical actions. Figure 1 shows an overview of this study.

2 RELATED WORKS

2.1 Research on Recognition of Tactical Actions

This study aims to estimate overlapped tactical actions. Similar to this study, there are studies that have investigated tactical action.

There are studies that recognize only specific tactical actions. First, several methods have been proposed for counter-attack recognition, which is also the estimated target of this study. Sigari et al. proposed a method to recognize counter-attacks based on camera motion (Sigari et al., 2015). Kobayashi et al. proposed a method to recognize counter-attacks using machine learning based on the position of agents in the RoboCup (Kobayashi et al., 2012). They proposed a model to recognize counter-attacks based on the following three characteristics: the collapse of the opponent's defense, the attack by a small number of players, and the movement of the ball into the opposition half. Dennis et al. used a Variational Auto-Encoder (VAE) to extract counter-attack features from unlabeled data. The authors extracted the characteristics of counter-attacks from unlabeled data using VAE (Fassmeyer et al., 2021). In addition, Bauer et al. proposed a method for recognizing and quantitatively evaluating the tactical action of counter-press (Bauer and Anzer, 2021). Forcher et al. proposed a quantitative evaluation method for defensive tactical actions (Forcher et al., 2022). Although there have been studies on the recognition of various tactical actions, recognition methods have not yet been established.

Suzuki et al. proposed a method to classify several tactical actions of a specific match situation through whole match (Suzuki et al., 2020). The classification was performed using deep learning with the opinions of experts as a teacher signal. However, it is believed that there are match situations that cannot be represented because they did not recognize tactical actions in overlapping. They considered exclusive relationships between the tactical actions performed by the two teams. They corrected the recognition results to account for exclusive relationships. Since we focus on estimating tactical actions in overlapping, we use a method that takes exclusive relationships into account when training the deep learning models.

2.2 Research Using the Temporal Changes of the Ball and Players Position

With the development and spread of tracking devices, tracking data has become readily available. This has led to research using the temporal changes of the ball and players' positions.

First, the temporal changes of the ball and players' position are used to analyze match situations. Lucey et al. examined a method for quantifying the possibility of chances by considering defensive positions and formations obtained from tracking data (Lucey et al., 2015). Kamiya et al. estimated the changes in match situations based on variables such as the ball position, front line position, and compactness (Kamiya et al., 2017).

In addition, indicators of team characteristics have been generated using tracking data. Lucey et al. proposed a method for capturing team characteristics and identifying teams from tracking data alone using multi-agent plan recognition (Lucey et al., 2012). Bialkowski et al. proposed a method for analyzing roles within a team by assigning each player soccer position (Bialkowski et al., 2014).

In this study, we also use the positions of the ball and players obtained from the broadcast videos for the estimation of tactical actions.

3 ESTIMATION OF OVERLAPPED TACTICAL ACTIONS

In this section, we describe a method for estimating overlapped tactical actions. We describe the details of the tactical actions we estimate in Section 3.1. We describe the estimation of each tactical action in Section 3.2. We describe the estimation of overlapped tactical actions in Section 3.3.

3.1 Description of Tactical Actions

This section describes tactical actions. There are eight types of tactical actions that we estimate. Four types of offensive tactical actions are long counter (lc), short counter (sc), opposition half possession (opp), and own half possession (own). Four types of defensive tactical actions are counter-press (cp), high press (hp), middle press (mp), and block (bl). FIFA states that there are 7 phases of play when the ball is in possession, attacking, and 9 phases of play when the ball is not in possession, defending (FIFA Training Centre, 2022). Among the 16 phases of play, 8 representative ones were selected as the tactical actions to be estimated.

The overlapped tactical actions have exclusive relationships. There are two types of exclusive relationships. First, there are tactical actions that cannot be performed by the two teams at the same time. Second, there are tactical actions that cannot be performed by the same team at the same time. In soccer, two teams play with one ball. It is impossible for the two teams to simultaneously perform a tactical action when it is in ball possession. It is also impossible for the same team to perform a tactical action when it is in ball possession and a tactical action when it is not in ball possession at the same time. In this study, the six tactical actions, excluding counter-attacks, are the tactical actions that cannot be performed by the two teams at the same time, as shown in the Table 1. Counter-attacks can overlap the execution times of the two teams. Long and short counters are tactical actions that cannot be performed by the same team at the same time, as shown in the Table 2. Other than the combination of long and short counters, the execution times of the two tactical actions can overlap in the transition.

3.2 Estimation of Each Tactical Action

In this section, we describe the estimation method for each tactical action. First, we organize the state variables used in this study. We describe the position of the ball and players and the sequence consisting of them. Next, we describe the teacher signal given to the sequence.

Table 1: Tactical actions that can be performed simultaneously by the two teams. Y represents tactical actions that can be performed by two teams simultaneously; N represents tactical actions that cannot be performed by two teams simultaneously.

		A Team							
		lc	sc	opp	own	cp	hp	mp	bl
B Team	lc	Y	Y	Y	Y	Y	Y	Y	Y
	sc	Y	Y	Y	Y	Y	Y	Y	Y
	opp	Y	Y	N	Y	Y	Y	Y	Y
	own	Y	Y	Y	N	Y	Y	Y	Y
	cp	Y	Y	Y	Y	N	Y	Y	Y
	hp	Y	Y	Y	Y	Y	N	Y	Y
	mp	Y	Y	Y	Y	Y	Y	N	Y
	bl	Y	Y	Y	Y	Y	Y	Y	N

Table 2: Tactical actions that can be performed simultaneously by the same team. Y represents tactical actions that a team can perform simultaneously; N represents tactical actions that a team cannot perform simultaneously.

		X Team							
		lc	sc	opp	own	cp	hp	mp	bl
X Team	lc	\	N	Y	Y	Y	Y	Y	Y
	sc	N	\	Y	Y	Y	Y	Y	Y
	opp	Y	Y	\	Y	Y	Y	Y	Y
	own	Y	Y	Y	\	Y	Y	Y	Y
	cp	Y	Y	Y	Y	\	Y	Y	Y
	hp	Y	Y	Y	Y	Y	\	Y	Y
	mp	Y	Y	Y	Y	Y	Y	\	Y
	bl	Y	Y	Y	Y	Y	Y	Y	\

The state variables used in this study are the positions of the ball and players on the field. The position of the ball is given in three-dimensional coordinates as shown in Fig. 2. The state variable is represented by the following equation. The position of the player is given by xy two-dimensional coordinates in Fig. 2.

$$\boldsymbol{O} = (O_x, O_y, O_z)^\top \tag{1}$$

$$\boldsymbol{A}[n] = (A[n]_x, A[n]_y)^\top \tag{2}$$

$$\boldsymbol{B}[n] = (B[n]_x, B[n]_y)^\top \tag{3}$$

$$\boldsymbol{P} = \{\boldsymbol{O}, \boldsymbol{A}[0], \ldots, \boldsymbol{A}[10], \boldsymbol{B}[0], \ldots, \boldsymbol{B}[10]\} \tag{4}$$

$$\boldsymbol{G} = \{\boldsymbol{P}_0, \ldots, \boldsymbol{P}_t, \ldots, \boldsymbol{P}_T\}^\top \; 0 \leq t \leq T \tag{5}$$

$$\boldsymbol{X}_t = \{\boldsymbol{P}_{t-l}, \ldots, \boldsymbol{P}_t\}^\top \tag{6}$$

The position of the ball at a given time is represented by Eq. (1). The position of the n-th player of A team is represented by Eq. (2). The position of the n-th player of B team is represented by Eq. (3). As

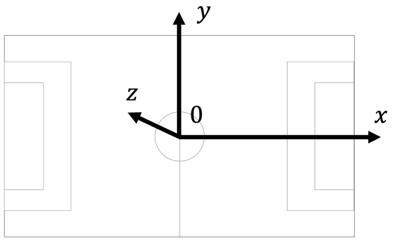

Figure 2: Direction of the 3D coordinates of the field.

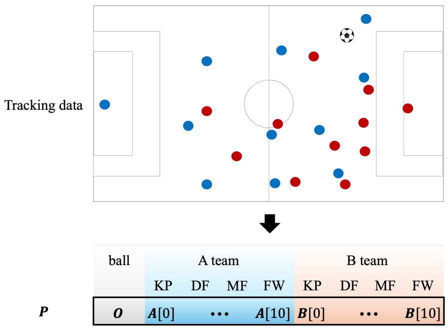

Figure 3: Sorting the positions of the ball and players in the tracking data by ball, A team keeper to forward, and B team keeper to forward.

in Eq. (4), $\boldsymbol{P}$ contains the positions of the ball, the players of A team, and the players of B team, in that order. The positions of the players of each team are stored in the order of keeper, defender, midfielder, and forward, as shown in Fig. 3. The position $\boldsymbol{P}$ at time t is denoted as $\boldsymbol{P}_t$. A match can be represented by the Eq. (5) using $\boldsymbol{P}_t$. To represent a specific match situation, a match is divided into sequences consisting of position at each time and past time. We denote by l the length of past time included in the sequence. The sequence $\boldsymbol{X}_t$ at a given time t can be expressed by Eq. (6).

Next, we describe the teacher signal. To train a deep learning model, it is necessary to provide the teacher signal $\boldsymbol{Y}_t$ along with the sequence $\boldsymbol{X}_t$ containing the positions of the ball and players. The teacher signal for a specific tactical action is denoted as $Y[tac action]_t$. The teacher signal indicates whether the tactical action is being performed at time t. We assign a value of 1 to the teacher signal $Y[tac\ action]_t$ when the tactical action is being performed and 0 when it is not. This is expressed by Eq. (7). We recognize tactical actions with simple definitions. We denote the start time of a tactical action as *start*, the end time as *end*, and the focus time as t. If the focus time t is between the start time *start* and the end time *end*, the tactical action is considered to be ongoing at time t.

$$Y[tac\ action]_t = \begin{cases} 0\ (t < start,\ end < t) \\ 1\ (start \leq t \leq end) \end{cases} \tag{7}$$

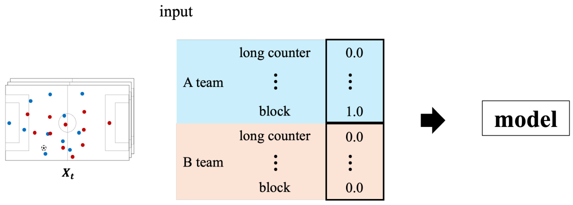

Figure 4: Teacher signal. We have the deep learning model to learn all tactical actions simultaneously.

3.3 Estimation of Overlapped Tactical Actions Considering Exclusive Relationships

We discuss the method to estimate overlapped tactical actions considering exclusive relationships. We have the deep learning model to learn all tactical actions simultaneously, as shown in Fig. 4. This allows for the estimation of overlapped tactical actions. When we train the deep learning model, we make the teacher signals of the tactical actions exclusive, which cannot be performed at the same time. This allows for the estimation considering exclusive relationships. The tactical actions that cannot be performed at the same time are based on Table 1 and 2.

4 IMPLEMENTATION

In this section, we describe how to implement the proposed method in Section 3. We describe the data used in this study in Section 4.1. We describe how to divide the data into sequences in Section 4.2. We describe the teacher signals for each sequence in Section 4.3. We describe the processes of training the model and the estimation in Section 4.4.

4.1 Data Format

In this study, two sets of data are used. The first is tracking data provided by SkillCorner. The second is event data provided by StatsBomb. We describe the tracking data. The three-dimensional coordinates of the ball and 22 players on the field are obtained at 10 fps. SkillCorner detects the ball and the players from the broadcast video and estimates their positions on the field. The positions of players not captured in the broadcast video are predicted using deep learning. Next, we describe the event data. Event data contains events such as passes, drives, shots, and goals, along with the time of occurrence and the position of the ball at that time. It also includes information such as whether the ball carrier was under pressure for each

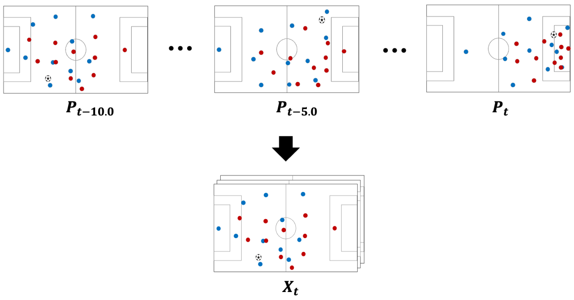

Figure 5: Sequence segmentation.

event. This information can be used to provide the teacher signals for tactical actions.

4.2 Sequence Segmentation

We describe the method for sequence segmentation. As shown in Eq. (6), the data is divided into sequences. The length of past time l is set to $l = 10.0$. As shown in Fig. 5, the sequence $\boldsymbol{X}_t$ at focus time t includes the position data from $t - 10.0$ to t. Since the fps is 10, the sequence contains 100 frames. The focus time t is slid 0.1 seconds at a time for segmentation. This allows the estimation results for each frame to be outputted.

Each frame contains a three-dimensional coordinate indicating the ball position, two-dimensional coordinates indicating the positions of 11 players from A team, and two-dimensional coordinates indicating the positions of 11 players from B team, as described by Eq. (4) and shown in Fig. 3. The players of each team are arranged in the order of goalkeeper, defender, midfielder, and forward. There is no specific rule regarding the order of players within the same position. Additionally, in Fig. 2, A team is set to attack in the positive direction, while B team is set to attack in the negative direction.

4.3 Acquiring Teacher Signals

We describe the method for obtaining the teacher signal. We handcraft a large number of tactical actions based on simple definitions. As shown in Fig. 6, the start and end times of each team's eight types of tactical actions are recognized from the event data using simple definitions. Next, the teacher signal at focus time t is obtained based on the start and end times of each tactical action. Finally, we reflect exclusive relationships of the tactical actions in the teacher signal.

First, we describe the recognition of tactical actions using simple definitions. We adapt the definitions of team style indicators published by the analysis company DataStadium. The definitions of each

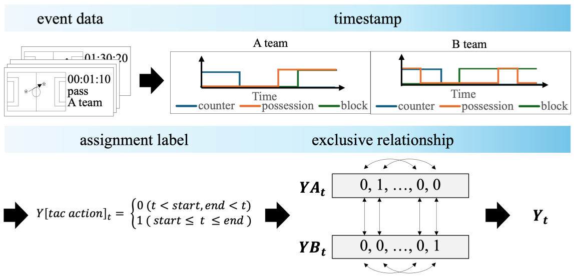

Figure 6: Assignment a teacher signal. We recognize start and end times of each tactical action from event data. We assign a teacher signal for the tactical action based on the start and end times.

tactical action are shown in Table 3. We have modified some of the definitions to simplify the recognition from event data. Based on these definitions, the start and end times of each tactical action are recognized from the event data through whole matches.

Next, we provide the teacher signal $\boldsymbol{Y}_t$ for each sequence $\boldsymbol{X}_t$. The teacher signal for A team's tactical actions is denoted as $\boldsymbol{YA}_t$. The teacher signal for B team's tactical actions is denoted as $\boldsymbol{YB}_t$. Since there are eight types of tactical actions, the teacher signal for one team consists of 8 values, as shown in Eqs. (8) and (9). To estimate the tactical actions of both teams simultaneously, the teacher signal $\boldsymbol{Y}_t$ is structured as shown in Eq. (10). As explained in Section 3.2, the teacher signal is either 0 or 1, as expressed by Eq. (7). The *start* and *end* are the start and end times recognized from the event data, and t refers to the focus time.

Finally, to consider exclusive relationships of tactical actions, we make the teacher signals of the tactical actions exclusive, which cannot be performed at the same time.

$$\begin{aligned}\boldsymbol{YA}_t = (&YA[lc]_t, YA[sc]_t, YA[opp]_t, YA[own]_t,\\ &YA[cp]_t, YA[hp]_t, YA[mp]_t, YA[bl]_t)\end{aligned} \tag{8}$$

$$\begin{aligned}\boldsymbol{YB}_t = (&YB[lc]_t, YB[sc]_t, YB[opp]_t, YB[own]_t,\\ &YB[cp]_t, YB[hp]_t, YB[mp]_t, YB[bl]_t)\end{aligned} \tag{9}$$

$$\boldsymbol{Y}_t = (\boldsymbol{YA}_t,\ \boldsymbol{YB}_t) \tag{10}$$

4.4 Training Model and Estimation

First, we describe the deep learning model used. Next, we describe the processes of training the model and the estimation.

The deep learning model used in this study is Long Short-Term Memory (LSTM). We use LSTM because it excels in supervised learning of time series data.

Table 3: Definition of tactical actions.

Tactical Action	Definition
long counter	Attacks that penetrate into the attacking third within 15 seconds after winning the ball in the defensive third
short counter	Attacks that penetrate into the attacking third within 10 seconds after winning the ball in the middle third or behind the attacking third
opposition half possession	Attacks with the ball possession for more than 20 seconds in opposition half
own half possession	Attacks with the ball possession for more than 20 seconds in own half
counter-press	Defenses that presses an opponent's ball carrier less than 5 seconds after the ball lost.
high press	Defenses that presses continuously against the play of the opponent's keeper or defenders
middle press	Defenses that presses continuously against the play of an opponent's midfielders or forwards
block	Defenses that do not press in own half when the opponent do the ball possession

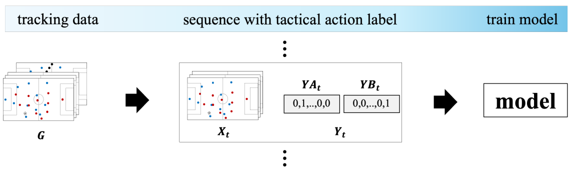

Figure 7: The process of training. We input sequences and teacher signals into a deep learning model.

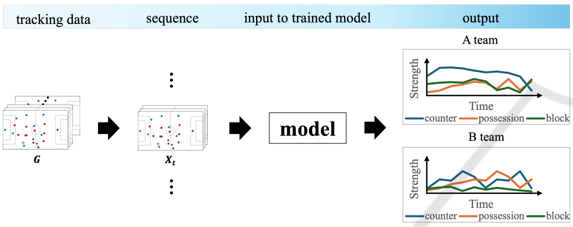

Figure 8: The process of estimation. We input a sequence into a trained model. The trained model outputs a estimation result for each tactical action of the two teams.

As shown in Fig. 7, the combinations of sequence $\boldsymbol{X}_t$ and teacher signal $\boldsymbol{Y}_t$ are input for training the LSTM model. Next, as shown in Fig. 8, when a sequence $\boldsymbol{X}_k$ is input to the trained LSTM model, the estimation result $\boldsymbol{Z}_k$ is output. The estimation results in $\boldsymbol{Z}_k$, similar to the teacher signal, consists of eight values for each team, as shown in Eqs. (11) and (12). The estimation result $\boldsymbol{Z}_k$ is composed of the estimation results of both teams as shown in Eq. (13). In certain tactical actions, the closer the estimation result is to 1, the higher the likelihood that the tactical action is being performed.

$$\boldsymbol{ZA}_k = (ZA[lc]_k, ZA[sc]_k, ZA[opp]_k, ZA[own]_k, \\ ZA[cp]_k, ZA[hp]_k, ZA[mp]_k, ZA[bl]_k) \tag{11}$$

$$\boldsymbol{ZB}_k = (ZB[lc]_k, ZB[sc]_k, ZB[opp]_k, ZB[own]_k, \\ ZB[cp]_k, ZB[hp]_k, ZB[mp]_k, ZB[bl]_k) \tag{12}$$

$$\boldsymbol{Z}_t = (\boldsymbol{ZA}_t,\ \boldsymbol{ZB}_t) \tag{13}$$

5 EXPERIMENT

In this section, we describe the estimation results of tactical actions using a trained model. First, we describe the dataset used for the experiments in Section 5.1. In Section 5.2, we present the estimation results for a particular scene. Finally, in Section 5.3, we describe the estimation results of tactical actions through whole matches as an ablation study.

5.1 Experimental Dataset

We describe the dataset used in this experiment. We use both a training dataset and a test dataset. First, we describe the training dataset. For the training dataset, we use tracking data from SkillCorner and event data from StatsBomb for 40 matches. All 40 matches are from the La Liga 2023-2024 season. From these 40 matches, we prepare 2500 sequences for each team's tactical actions, except for counter-pressing, which has 2000 sequences. A total of 39,000 sequences and their corresponding teacher signals are used to train the LSTM model. Next, we describe the test dataset. We prepare a test dataset of 10 matches, all from the La Liga 2023-2024 season, but not included in the training dataset. After training the LSTM model with the training dataset, we input the test dataset into the trained LSTM model to verify the estimation results. Finally, we describe a particular scene that we evaluate qualitatively. We use the scene with overlapped tactical actions. The scene is a one-minute period from 43:30 to 44:30 in the second half of the Real Madrid vs. Barcelona match. The scene is illustrated using images that reflect the positions of the ball and the players on the field of the CG environment based on the tracking data, as shown in Figure 9. In this scene, the players are switching between offense and defense, and overlapped tactical actions are occurring in between.

5.2 Estimation Results of Overlapped Tactical Actions

In this section, we describe the estimation results of the situation with overlapped tactical actions using the trained LSTM model. From this experiment, we confirmed that the trained model can estimate overlapped tactical actions considering exclusive relationships. As the results show, we succeeded in expressing the tactical actions performed at a given time in

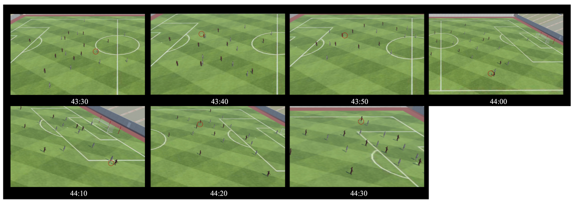

Figure 9: The scene with overlapped tactical actions. The scene is a one-minute period from 43:30 to 44:30 in the second half of the Real Madrid vs. Barcelona match

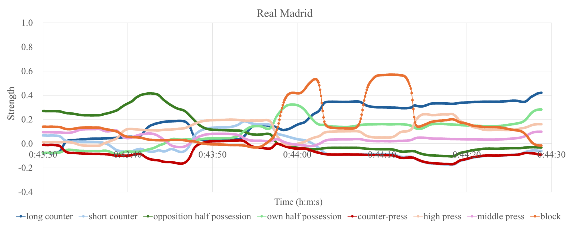

Figure 10: Estimation results of tactical actions of Real Madrid.

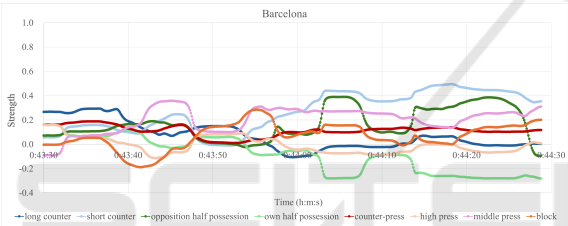

Figure 11: Estimation results of tactical actions of Barcelona.

terms of the strength of several tactical actions. Currently, the dataset we handcrafted does not represent the strength of tactical actions as time-varying. To verify the performance for continuous estimation, we need to expand the dataset for this purpose in the future.

Figure 10 shows the estimation results for Real Madrid's tactical actions. Figure 11 shows the estimation results for Barcelona's tactical actions. The trained model estimated that Real Madrid was attacking with opposition half possession from 43:30 to 43:50. At the same time, it estimated a high possibility that Barcelona was performing long counter and middle press. The estimation results changed as the attack and defense roles reversed at around 43:50. The model estimated a high possibility that Real Madrid was performing block and long counter while Barcelona was performing short counter and opposition half possession. In the actual match, around 44:30, Real Madrid regained the ball and attempted a counter-attack, but it was stopped by a foul. The trained model estimated this as a long counter. The model did not estimate the likely to be performed two tactical actions that cannot be performed at the same

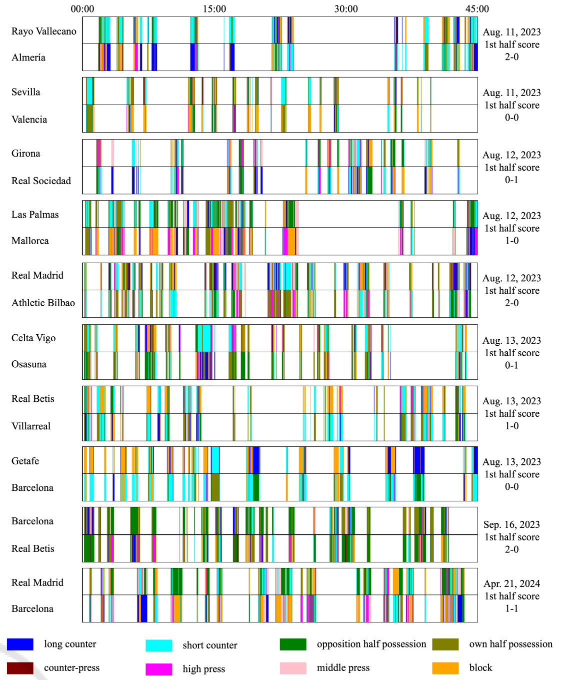

Figure 12: Estimation results for the first 45 minutes of the 10 games. This graph shows the tactical actions estimated as most likely to be performed by each team at each time.

time based on Table 1 and 2. We believe that the estimation of overlapped tactical actions considering exclusive relationships was achieved.

5.3 Ablation Study

In this section, we describe the estimation results of tactical actions through whole matches using a test dataset of 10 matches. From this experiment, we confirmed the validity of the estimation results. Additionally, we demonstrated the potential to represent match situations through the tactical actions performed by the two teams. Figure 12 shows the tactical actions that were estimated as most likely to be performed by each team at each time. We present the estimation results for the first 45 minutes of each match. Tactical actions are estimated only in situations where play continues for more than 10 seconds, as each sequence requires 10 seconds of position data. Looking at Fig. 12 as a whole, the team with the superior score often performs offensive tactical actions most of the time. Conversely, inferior teams often perform defensive tactical actions most of the time. By comparing the last three matches, we can confirm the correlation between estimation results and the objective match content using Barcelona as an example. In the match against Real Betis, Barcelona had an advantage in terms of score. The estimation results show a dominance of possession attacks. In the match against Real Madrid, the team performed defensive tactical actions most of the time, indicating a struggle against a strong opponent. In the match against Getafe, although offensive tactical actions were predominant, the score was 0-0. The estimation results predict that Barcelona struggled to convert their attacks into goals. In fact, Barcelona's expected goals (xG) in this match were significantly higher than Getafe's, but they failed to score. From these observations, we confirmed that the estimation results align with the objective match content. Furthermore, by representing match situations using the combination of tactical actions of the two teams, we demonstrated the potential to predict match situations.

6 CONCLUSION

In this study, we propose a new method for estimating overlapped tactical actions from soccer match videos. We enable the estimation of overlapped tactical actions considering exclusive relationships. We achieve this by having the deep learning model learn all tactical actions simultaneously. We handcraft a large number of tactical actions based on simple definitions to estimate tactical actions using deep learning. We estimate tactical actions in deep learning by focusing on the temporal changes of the positional relationships among 22 players and between the ball and the players.

We validated the estimation of overlapped tactical actions considering exclusive relationships by using 10 matches. We succeeded in expressing the tactical actions performed at a given time in terms of the strength of several tactical actions. In the ablation study, we confirmed the validity of the estimation results.

ACKNOWLEDGEMENTS

Part of this research is supported by JSPS KAKENHI Grand Number 21H03476 and 21H05300.

REFERENCES

Barnerat, T., Crevoisier, J., Hoek, F., Redon, P., and Ritschard, M. (2000). *FIFA Coaching*. Druckerei Feldegg AG.

Bauer, P. and Anzer, G. (2021). Data-driven detection of counterpressing in professional football: A supervised machine learning task based on synchronized positional and event data with expert-based feature extraction. *Data Mining and Knowledge Discovery*, 35.

Bialkowski, A., Lucey, P., Carr, P., Yue, Y., Sridharan, S., and Matthews, I. (2014). Large-scale analysis of soccer matches using spatiotemporal tracking data. In *2014 IEEE International Conference on Data Mining*, pages 725–730.

Fassmeyer, D., Anzer, G., Bauer, P., and Brefeld, U. (2021). Toward automatically labeling situations in soccer. *Frontiers in Sports and Active Living*, 3.

FIFA Training Centre (2022). EFI Explanation Documents (PDF) — fifatrainingcentre.com. https://www.fifatrainingcentre.com/en/fwc2022/efi-metrics/efi-metrics-pdfs.php. [accessed on 13th July 2024].

Forcher, L., Altmann, S., Forcher, L., Jekauc, D., and Kempe, M. (2022). The use of player tracking data to analyze defensive play in professional soccer - a scoping review. *International journal of Sports Science & Coaching*, 17:174795412210757.

Kamiya, K., Nakanishi, W., and Izumi, Y. (2017). Tracking data to extract changes in football game situation. *Proceedings of the Institute of Statistical Mathematics*, 65.

Kobayashi, Y., Kawamura, H., and Suzuki, K. (2012). Counter attack detection with machine learning from log files of robocup simulation. In *The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligence Systems*, pages 1821–1826.

Lucey, P., Bialkowski, A., Carr, P., Foote, E., and Matthews, I. (2012). Characterizing multi-agent team behavior from partial team tracings: Evidence from the english premier league. In *The AAAI Conference on Artificial Intelligence*, pages 1387–1393.

Lucey, P., Bialkowski, A., Monfort, M., Carr, P., and Matthews, I. (2015). "quality vs quantity": Improved shot prediction in soccer using strategic features from spatiotemporal data. In *The 8th Annual MIT Sloan Sports Analytics Conference*, pages 1–9.

Sigari, M.-H., Soltanian-Zadeh, H., Kiani, V., and Pourreza, A.-R. (2015). Counterattack detection in broadcast soccer videos using camera motion estimation. In *2015 The International Symposium on Artificial Intelligence and Signal Processing (AISP)*, pages 101–106.

Suzuki, G., Takahashi, S., Ogawa, T., and Haseyama, M. (2020). An estimation method of candidate region for superimposing information based on gaze tracking data in soccer videos. In *2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, pages 1–2.

Wang, Q., Zhu, H., Hu, W., Shen, Z., and Yao, Y. (2015). Discerning tactical patterns for professional soccer teams: an enhanced topic model with applications. In *The 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2197–2206.