RESEARCH ARTICLE

Multi-Agent Deep-Learning Based Comparative Analysis of Team Sport Trajectories

ZHANG ZIYI[1], RORY BUNKER [ID][1], KAZUYA TAKEDA [ID][1], (Senior Member, IEEE), AND KEISUKE FUJII [ID][1,2,3], (Member, IEEE)

[1]Graduate School of Informatics, Nagoya University, Nagoya, Aichi 464-8601, Japan
[2]RIKEN Center for Advanced Intelligence Project, Osaka 103-0027, Japan
[3]JST PRESTO, 7 Gobancho, Chiyoda, Tokyo 332-0012, Japan

Corresponding author: Keisuke Fujii (fujii@i.nagoya-u.ac.jp)

ABSTRACT Computational analysis of multi-agent trajectories is a fundamental issue in the study of real-world biological agents. For trajectory analysis, combining movement data with labels (e.g., whether a team scores in a ball game) can provide additional insights compared to relying only on trajectory data. However, existing deep-learning-based methods consider only single-agent animal trajectories, and cannot be directly applied to multi-agent trajectories in sports. In this paper, a comparative analysis method to analyze multi-agent trajectories in ball games is proposed. A neural network approach based on an attention mechanism using multi-agent motion characteristics (e.g., the distances between agents and objects) as the input is adopted, which is designed to detect distinct segments in trajectories of given classes. This enables us to understand differences between classes by highlighting segmented trajectories and which variables correlate with the given labels. The effectiveness of our approach was verified by comparing various baselines with effective/ineffective attack labels and goal/non-goal labels using different sizes of the dataset. The effectiveness of our method is also demonstrated by analyzing the attacking plays in an NBA dataset.

INDEX TERMS Machine learning, trajectory analysis, interpretability, multi-agent systems.

I. INTRODUCTION

The development of measurement technologies, e.g., global/local positioning and camera-based systems has enabled the computational analysis of multiple living agents in the real world. Recent progress has led to an improved comprehension of the underlying principles governing real-world multi-agent behaviors, which is a fundamental issue across several scientific and engineering domains. When the elements of real-world multi-agent systems are not physically linked, their underlying rules are often poorly understood [1], [2]. To gain insight into multi-agent movements, mathematical models that rely on basic rules can be used. For instance, social force models [3], which posit a force acting among individuals, have been widely used to study pedestrian behaviors. These models, under certain assumptions, have also

The associate editor coordinating the review of this manuscript and approving it for publication was Rongbo Zhu [ID].

been applied to more complicated behaviors such as those observed in team sports [4], [5], [6]. However, the complexity of higher-order social interactions, cognition, and body dynamics makes it mathematically challenging to model the general multi-agent behaviors of living organisms such as in team sports [7]. Consequently, a model-free (or equation-free) and data-driven approach is necessary to improve our understanding of these behaviors [2], [8].

Data-driven modeling is a powerful tool that can extract information and make predictions from complex real-world data. Machine learning has actively studied the learning process of models with complex, nonlinear structures such as neural networks [9]. Although these models can offer higher expressiveness and predictive performance, their results can be challenging to interpret, creating a trade-off between interpretability and expressiveness (or predictability). This challenge is particularly important for practical applications in actual sports games, where coaches and players need
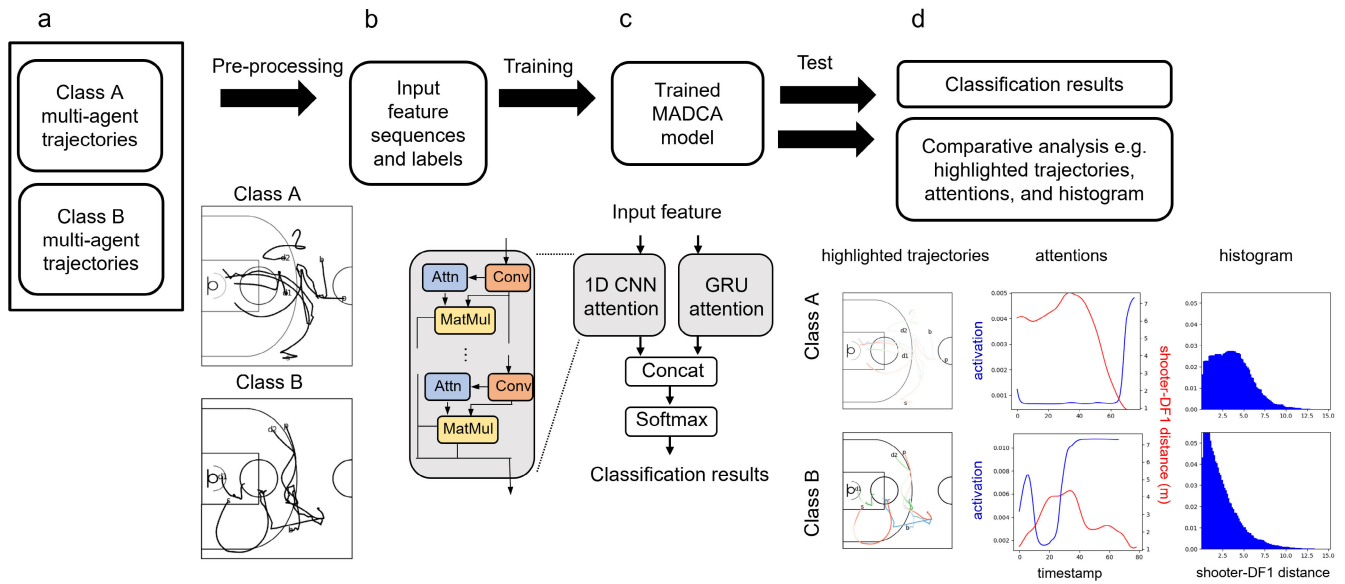
FIGURE 1. Our framework of multi-agent trajectory comparative analysis using MADCA. (a) Multi-agent trajectories of classes A and B are given. In this example, the initial positions of the ball, shooter (including a ball lost case), passer, and defenders 1 and 2 are presented as b, s, p, d1, and d2, respectively. (b) Input feature sequences are computed by pre-processing. (c) A MADCA model called MADCA-net comprises a 1D convolutional neural network (CNN) and gated recurrent unit (GRU) with attention mechanism and outputs classification results. (d) The main outputs are classification results, highlighted trajectories (left), attentions for each layer and time (middle), and attended features (right). For details, see Sections III-B and III-E.

information about why a goal was scored and the characteristics observed in subsequent plays.

Currently, the trajectory data of players and the ball in professional sports (e.g., basketball or soccer) can be accessed. For trajectory analysis, combining with labels (e.g., good or bad attacks in ball games) can provide insights compared to only trajectory data. There have been many approaches for supervised and unsupervised learning of multi-agent trajectory data (see Section II). Compared with previous approaches, the analysis of multiple player trajectories by highlighting the difference between labels to understand multi-agent behaviors is focused on. In a different research field, that of animal trajectory analysis, comparing two data classes to obtain useful insights is referred to as comparative analysis using DeepHL [10]. However, this deep-learning-based method used only single-agent animal trajectories; multi-agent motion characteristics (e.g., the distances between agents) were not considered.

In this paper, a comparative analysis method to analyze multi-agent trajectories in a ball game is proposed, which we call Multi-Agent Deep-learning based Comparative Analysis (MADCA, Fig 1). A neural network approach based on an attention mechanismis adopted, which is the combination of convolutional neural network (CNN) and an recurrent neural network (RNN), to detect distinct segments in trajectories of given classes (Fig 1c). This method enables us to understand the differences between classes by highlighting segmented trajectories and identifying which variables correlate with the labels (Fig 1d). Our approach was verified by comparing various ablated models and demonstrated its effectiveness

through use cases that analyze the difference between effective and ineffective attacks in US National Basketball Association (NBA) games. For example, based on the correlation between attention values and the handcrafted features, the distance between the shooter and the shooter defender was selected and the histograms of the selected feature were clearly different between the effective and ineffective attack classes.

The main contributions of this paper are as follows:

(i) MADCA is proposed, a comparative analysis method to analyze multi-agent trajectories in ball games, the goal of which is to understand the differences between classes by highlighting segmented trajectories and which variables correlate with the labels.

(ii) A neural network approach based on an attention mechanism is adopted that uses multi-agent motion characteristics (e.g., the distances between agents and objects) as input and detects distinct segments in trajectories of given classes.

(iii) Our approach is verified by comparing various ablated models with effective/ineffective attack labels and goal/non-goal labels, using different sizes of the dataset. We also demonstrate the effectiveness of our method through use cases that analyze the difference between effective and ineffective attacks in the NBA dataset.

The remainder of this paper is organized as follows. First, an overview of related work is presented in Section II. Then, our method is described in Section III. The experimental results are presented and discussed in Sections IV and V, respectively.

II. RELATED WORK

Traditional methods without machine learning in various fields typically rely on researchers' experience and established theories to evaluate the characteristics of multi-agent behaviors. For example, researchers have calculated the distances and relative phases of two athletes (e.g., [11], [7], [12]), speeds of movements (e.g., [13]), frequencies and angles of actions (e.g., shots [14] and passes [15], [16], [17]), as well as their representative values (e.g., average and maximum values). Although this quantitative approach is powerful, easy to interpret, and applicable to small datasets in specific sports, it may not be flexible enough to represent cooperative/competitive interactions in detail. To address this limitation, various data-driven methods, including machine learning techniques, have been developed to extract features and rules automatically.

Learning-based approaches that utilize positional data of players can be broadly categorized into unsupervised and supervised learning approaches. Unsupervised learning methods allow algorithms to infer functions that describe hidden structures from unlabeled data, which can be useful for knowledge discovery from data without clear hypotheses. Typical unsupervised methods in sports trajectory analysis include dimensionality reduction [1], [8], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27] and clustering [28], [29], [30], [31], [32], [33], [34]. However, the lack of objective variables in unsupervised learning may make it difficult to interpret the results and verify the methods when complex models are used.

Supervised learning methods, which infer a function from labeled training data, can be useful for classification and regression. A simple approach is to input static features for classification or regression (e.g., [35], [36], [37], [38], [39], [40], [41], [42]) using, e.g., linear discriminant analysis (LDA), logistic regression, or support vector machine (SVM) with hand-crafted static features. However, complex multi-agent behaviors often require reflecting the time-series structure of data, which can be achieved by combining dynamic features with unsupervised learning or neural networks [1], [8], [23], [27], [43], [44], [45]. Although neural network approaches are flexible (e.g., trajectory prediction [46], [47], [48], [49], [50], [51]), they sometimes lack interpretability. To obtain interpretable spatial representations, several approaches have been developed such as using matrix [52] and tensor [53], [54] factor models, and Poisson point processes [55] that focus on on-ball behaviors. Compared with these approaches, the analysis of the trajectories of multiple players to understand multi-agent behaviors is focused on.

III. METHODS

In this section, the dataset is first described, then our machine learning model, preprocessing, and analysis procedures.

A. DATASET

A basketball dataset from the NBA 2015-2016 season, preprocessed by the STATS SportVU system (Northbrook, IL, USA) was used, which contains the positional data of players and the ball (at a frequency of 25 frames per second). 600 games from the dataset was chosen because we deemed that this represented sufficient data. The positional data contained the (x,y) positions of each player on the court, and the (x,y,z) coordinates of the ball. The dataset was divided beforehand into attack segments from the start of the attack (ball possession of the team or already divided segmentation in the raw data) to the transition to the next attack. The end of the attack segment is defined as being when a shot is made or the ball is lost (known as a turnover in basketball). The data contained a total of 45,307 attacks, sub-sampled at 10 Hz, i.e., the time between each point coordinate is always 0.1 seconds. In our dataset, there were 18,021 shot successes, 27,286 shot failures (including 7131 turnovers), 22,159 effective attacks, and 23,148 ineffective attacks (the definitions of effective/ineffective attacks are described in Section III-C). The probabilities of scoring, given the attack was effective and ineffective, were 0.466 and 0.333, respectively, which indicates that the effective attack indicator is valid in terms of scoring on average. However, in a strict sense, scoring and effective attack are different (for further detail, see Section III-C).

In our analysis, the trajectories of five agents (Fig 1a) is considered. These five agents comprise the ball and four players: the shooter (or the last player who was on the ball, including a ball lost case), the defender of the shooter at the last frame (called DF1), the last passer to the shooter (called passer), and the defender of the last passer at the last frame (called DF2). These agents were selected because the verification of our approach is focused on and all trajectories may be too diverse for the model to learn a good representation for highlighting the differences between the two labels. In general, this is a multi-agent role assignment problem for an unsorted diverse dataset (see e.g., [31]). This problem is avoided by using only predetermined roles about four players and the ball. It is considered that it is more difficult to determine the roles in a fixed manner as the number of players increases, and fewer players may be less informative in this analysis. Then the interval from the ball-receiving time of the passer to the end of the above attack segment was analyzed.

B. PROPOSED MODEL

The proposed method is designed to help understand the differences between classes of multi-agent trajectories in sports by highlighting segmented trajectories and which variables correlate with the labels. To this end, a comparative analysis method is proposed to analyze multi-agent trajectories called MADCA, which extends the single-agent trajectory DeepHL framework [10] to the multi-agent trajectory problem.

Similar to [10], our method assumes that there are two classes of trajectory data with different properties such that each trajectory belongs either to class A or B (e.g., scored or not). Here, the neural networks for MADCA is explained, which are shown in Fig. 1c. The pipeline of MADCA is briefly introduced such that:

(i) Our network (hereafter called MADCA-Net) is first trained on the trajectory data from two classes, which is the combination of CNN and RNN.

(ii) The attention mechanism in MADCA-Net computes the attention value of each time stamp of the trajectory.

(iii) After obtaining the attention, distinctive parts of the trajectories in two classes are highlighted using the attention in a particular layer. To find such a layer (hereafter referred to as a "distinctive layer"), the score is computed for each layer using the attention value.

(iv) MADCA also extracts the highlighted segments with handcrafted features, which is based on the correlation between the attention values and the handcrafted features (Fig. 1d middle).

Next, the input of the model is described. The input is a time series of features, an $l_{MAX} \times N_f$ matrix, where l_{MAX} is the maximum length of the input trajectories and N_f is the dimensionality of the features. The features in DeepHL [10] include basic features used in locomotion analysis (e.g., position and speed). In this study, multi-agent motion features such as the distances between the K agents and an object (e.g., a goal called a ring in basketball) are used. For further details of the input features, see Subsection III-C. Since the lengths of the trajectories are not identical to each other, the missing elements are masked when training the network.

MADCA-Net classifies a trajectory into either class and to output the segments of a trajectory to which the distinctive layer pays attention. Figure 1c shows the architecture of MADCA-Net, comprising four stacks of 1D convolutional layers and a gated recurrent unit (GRU) [56] layer, which is one of RNN architectures. As used in DeepHL [10], the 1D convolutional layers (the orange-colored blocks in Fig. 1c) extract short-term features. To extract features at different levels of scale, in each 1D convolutional layer, we extract features using a kernel size or filter width F_t, which are $3\%, 6\%, 9\%,$ and 12% of l_{MAX} in the four convolutional stacks. A step size or stride of one sample is used in terms of the time axis. Padding is employed to ensure that the length of the outputs of a given layer are corresponded with that of the inputs to the layer. The convolutional stacks are constructed to extract features at different levels of scale by utilizing different filter sizes across the different stacks.

In contrast, the GRU layers tend to extract features reflecting long-term dependencies (the configuration of the attention mechanism is the same as in the 1D convolutional layers). Compared with DeepHL [10], which uses long short-term memory (LSTM), we used a GRU, which has a smaller number of parameters than LSTM, since in preliminary experiments, the original DeepHL models took a long time to train. In addition, because we believe last-moment

information (e.g., shot) will be important, and the different levels of abstraction may not be important for multi-agent trajectory data in sports, a simple two-layer GRU is implemented rather than four stacks of LSTMs with different layer sizes [10].

In order to identify which segments of the trajectories are significant for each layer, an attention mechanism [57] is incorporated into the model as illustrated in Fig. 1c. The output of each 1D convolutional/GRU layer for an input trajectory is used to calculate the attention vector such that

$$a = \mathrm{softmax}\left(\tanh(W_a Z^\top + b_a)\right), \qquad (1)$$

where $a \in \mathbb{R}^{1 \times l_{\mathrm{MAX}}}$. This shows the importance of each time stamp in the trajectory and is also used to highlight parts of the trajectory. An attention vector has the same length as the trajectory, where l_{MAX} is the maximum length of the input trajectories. Matrix $Z \in \mathbb{R}^{l_{\mathrm{MAX}} \times N}$ is an output of the 1D convolutional/GRU layer, where N is the number of nodes in the convolutional/GRU layer. $W_a \in \mathbb{R}^{1 \times N}$ and $b_a \in \mathbb{R}^{1 \times l_{\mathrm{MAX}}}$ are the weight matrix and its bias, respectively. The softmax function ensures that the sum of all output values is equal to one, while the tanh function constrains the output value of its input to a range between -1 and 1. The attention mechanism is implemented as a neural network in MADCA-Net, specifically using layer-wise attention as indicated by the aqua-colored blocks in Fig. 1c. The attention vector a is multiplied by the outputs of the 1D convolutional/GRU layer using matrix multiplication (MatMul), as shown by the khaki-colored blocks in Fig. 1c. The outputs of all layers are concatenated and then used to estimate the class (Class A or B) in the final layer of MADCA-Net, which is a densely connected output layer using the softmax function, as shown in Fig. 1c. As mentioned above, the model is designed to calculate attention information at different levels of scale using 1D convolutional/GRU layers.

It is noteworthy that the parameters in Eq. (1) for each layer, including W_a, b_a, and the parameters in the convolutional and GRU layers, are estimated in the training. The tanh activation function is introduced into Eq. (1) to smooth the output attention. Without the tanh function, if an outlying large value is included in $W_a Z^\top + b_a$ at time t, attention values other than those at time t become extremely small. This causes only one data point to be colored in red when visualizing a trajectory using such attention values, making it difficult to identify important segments.

For processing the layers in MADCA-Net, a scoring system is employed, similar to that used in DeepHL [10], using the following equation:

$$s(A_{i,C_A}, A_{i,C_B}) = s_{\mathrm{fc}}(A_{i,C_A}, A_{i,C_B}) + s_{\mathrm{it}}(A_{i,C_A}, A_{i,C_B}), \quad (2)$$

where A_{i,C_A} and A_{i,C_B} with the ith layer are sets of attention vectors calculated from trajectories belonging to class A and B, respectively. Since an attention vector from a distinctive layer is expected to exhibit high values within a restricted range of segments, $s_{\mathrm{fc}}(A_{i,C_A}, A_{i,C_B})$ computes the average

variance of the attention values, which is normalized based on the average trajectory length, as follows:

$$
\begin{aligned}
&s_{\mathrm{fc}}(A_{i,C_A}, A_{i,C_B}) \\
&= \sqrt{\frac{1}{|A_{i,C_A} \cup A_{i,C_B}| \cdot l(A_{i,C_A} \cup A_{i,C_B})} \sum_{a \in A_{i,C_A} \cup A_{i,C_B}} \mathrm{Var}(a),}
\end{aligned}
\tag{3}
$$

where $\mathrm{Var}(\cdot)$ computes the variance and $l(\cdot)$ computes the average trajectory length. Using $l(A_{i,C_A} \cup A_{i,C_B})$, the computed variance is normalized. To prevent a larger variance for longer trajectories because the softmax function in Eq. (1) ensures that all values sum to one, the average variance is normalized using the average trajectory length.

Additionally, to assess the difference in attention value distributions between the two classes, the score $s_{\mathrm{it}}(A_{i,C_A}, A_{i,C_B})$ is computed as follows:

$$
s_{\mathrm{it}}(A_{i,C_A}, A_{i,C_B}) = (1 - \mathrm{Intersect}(h(A_{i,C_A}), h(A_{i,C_B}))). \tag{4}
$$

Based on DeepHL [10], $h(\cdot)$ computes a normalized histogram of attention values with 200 bins, and Intersect$(\cdot, \cdot)$ computes the area of overlap between the two histograms. Specifically, the computation is given by the equation:

$$
\mathrm{Intersect}(H_1, H_2) = \sum_i \min(H_1(i), H_2(i)), \tag{5}
$$

where $H_1(i)$ denotes the normalized frequency of the ith bin of histogram H_1.

C. PROCESSING PROCEDURE

Here, the steps taken in this study are described to compute the input features and target labels for MADCA-net. For the input features, the variables used in DeepHL [10] are first computed: position, speed, and distance from the initial location for each agent. The location data have two dimensions, while the speed (the norm of two-dimensional velocity) and distance from the initial location have one dimension for each agent. In this study, as multi-agent features, various distances among the agents and the ring (a fixed object) were added: shooter-DF1, shooter-DF2, passer-DF1, passer-DF2, shooter-ring, passer-ring, and ball-ring. Furthermore, the moving average, and moving variance of the above variables as per DeepHL [10] are computed.

Next, the steps are described to compute target labels. Two types of labels are considered: goal/no-goal and effective/ineffective attacks. The two MADCA-Nets were trained separately using the two types of labels. The goal/no-goal label can be straightforwardly defined based on the results of the attacks. However, since goal predictions are difficult in general (e.g., [8], [27]), another label was defined to be based on whether or not a particular play was an "effective attack," rather than whether a goal was scored in a particular play.

The tactics and strategy of a coach and team may be most influential up until the point at which there is a good scoring opportunity to make a shot, and it is then the skills/form of the individual player that determines whether this opportunity is

actually converted into a goal. A good scoring opportunity in basketball is considered to be a shot being attempted in a context in which there is a high expected probability of scoring, based on historical attempted and successful shots. Therefore, an interpretable and simple indicator is computed from available statistics (i.e., based on the frequency) to evaluate whether a player makes an effective shot attempt, rather than using a label based on whether a goal was scored or a learning-based score prediction model.

From the available NBA statistics, two basic factors were focused on for effective attacks at an individual player level: the shot zone on the court and the distance between a shooter and the nearest defender. This is based on the previous work [51] that only uses shooter and DF1 information (not all four players). These two factors are considered important for basketball successful shot prediction [7], [8], [27]. In the NBA advanced stats [58], Probabilities of successful shots attempted in each zone and distances for each player can be accessed. The shot zones are separated into four areas: the restricted area, in-the-paint, mid-range, and 3-point area. The restricted area is defined as the area within a radius of 2.44 m (the distance between the side of the rectangle and the ring) from the ring. The in-the-paint area is defined as the area within a radius of 5.46 m (the distance between the ring and the farthest vertex of the rectangle) from the ring. The three-point area is defined as the area that is outside of the 3-point line. The mid-range area is the remaining area. The shooter's distance from the nearest defender is categorized into four ranges: 0-2 feet, 2-4 feet, 4-6 feet, and more than 6 feet.

An effective attack is defined as one that meets the following criteria:

- The shooter's position in the restricted area is effective at any distance (because a defender is often located near the shooter).
- The shooter's position in the paint and mid-range is effective at a distance of 6 feet or more from the nearest defender (this range is regarded as "open" in the NBA advanced stats.
- The shooter's position in the 3-point area is effective when a player with a shot success probability of at least 0.35 attempts a shot at a distance of 6 feet or more from the nearest defender (because some players do not shoot tactically).

Based on the statistics in the 2014/2015 season and the tracking data, The probabilities of successful shots were computed for each zone and the distances for each player. The probability of the player who attempted shots less than 10 times was computed as the probability of the player that is of the same position (i.e., guard, forward, center, guard/forward, and forward/center based on their registration in the NBA 2014/2015 season). It should be noted that certain characteristics of a good shot can differ depending on the court location and context, for example, for 2-point and 3-point shots. Note that, unfortunately, those for only two areas (the 2-point and 3-point areas), with four distance categories could be
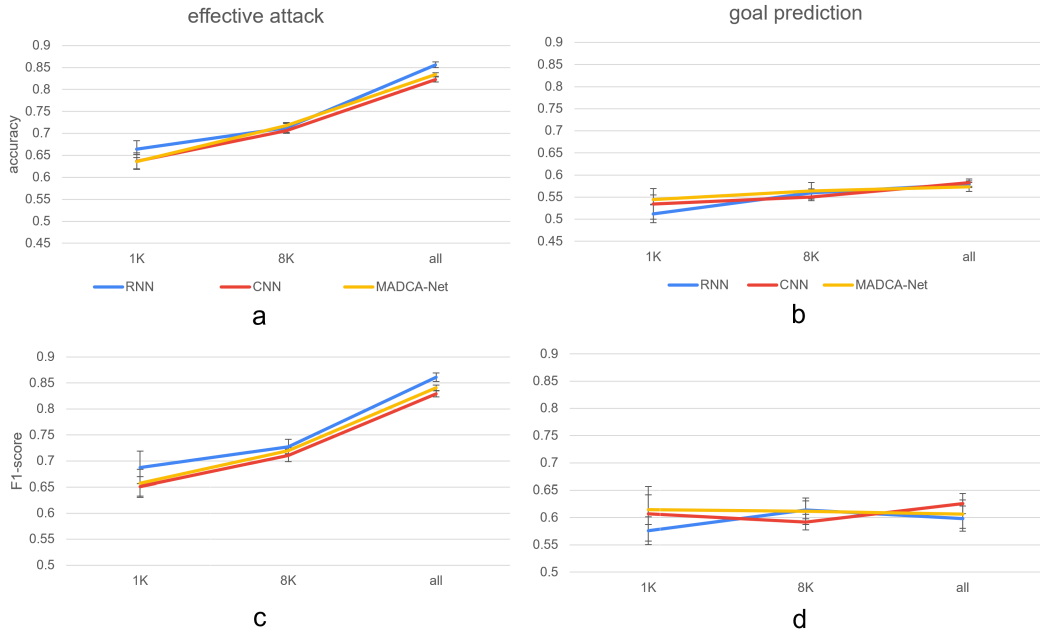
FIGURE 2. Prediction performances of all models. The effective/ineffective attack prediction and goal/non-goal prediction accuracies (a,b) and F1-scores (c,d) are shown.

accessed. Thus, the shot success probabilities in the restricted, in-the-paint, and mid-range areas were computed using those from the 2-point area.

D. TRAINING

MADCA-Net was trained on 80% of the trajectories, which were randomly selected, to minimize the binary classification error on the training data, employing backpropagation based on the Adam optimizer. Then, the trained MADCA-Net was tested using the remaining 20% of trajectories to compute the classification accuracy. All models were trained for 50 epochs, and there were 128 neurons in each convolutional/GRU layer. To reduce an overfitting, dropout was used with a rate of 0.5.

E. ANALYSIS

In our analysis, our methods were first validated in terms of classification performance. Then, example visual analyses were shown using our framework. Lastly, analyses of team performances were demonstrated.

To validate our methods in terms of classification performance, we used accuracy and F1-score metrics. Although the accuracy evaluates both true positives and negatives, the F1-score evaluates whether the true positives can be classified without considering the true negatives. The F1-score is expressed as

$$\text{F1score} = \frac{2 \times \text{Precision} \times \text{Recall}}{(\text{Precision} + \text{Recall})}, \quad (6)$$

where the Recall is defined as the ratio of the sum of true positives and true negatives to the number of true positives (the true-positive rate), and the Precision is defined as the ratio of the sum of true positives and true negatives to

false positives. When comparing our full model, CNN-RNN (MADCA), to two ablated models, separating 1D CNN and GRU (RNN) models was considered. Our approach was verified by comparing the models with effective/ineffective attack labels and goal/non-goal labels, using different dataset sizes (1,024, 8,192, and all 45,307 samples). It is speculated that 1K is a minimum size of the training and 8K is roughly an intermediate size between the minimum and full sample sizes in a log scale. Note that it is not obvious that more data provides a better result in this dataset and these tasks if a classification task is inherently difficult or some aspect of the model or input features is wrong. To examine these possibilities, we verified our approach using different dataset sizes. With 5 different random seeds when splitting the data into training and testing sets, the mean and standard deviation of the classification performances were evaluated.

To understand the meaning of the highlights, as per DeepHL [10], the Pearson correlation coefficients between the attention values of each layer and handcrafted features was computed as shown in Fig. 1d middle. Based on the correlation coefficients, the highlighted trajectories (Fig. 1d left) were ploted. For feature analysis, the differences between the distributions of each handcrafted feature for two classes within the highlighted segments were computed [10] as follows:

$$\text{diff}(A_{i,C_A}, F_{j,C_A}, A_{i,C_B}, F_{j,C_B})$$
$$= 1 - \text{Intersect}\left(h(m(A_{i,C_A}, F_{j,C_A})), h(m(A_{i,C_B}, F_{j,C_B}))\right), \quad (7)$$

where F_{j,C_A} is a set of time series of the jth handcrafted feature, calculated from trajectories belonging to class A. In addition, $m(\cdot, \cdot)$ is a masking function that extracts feature
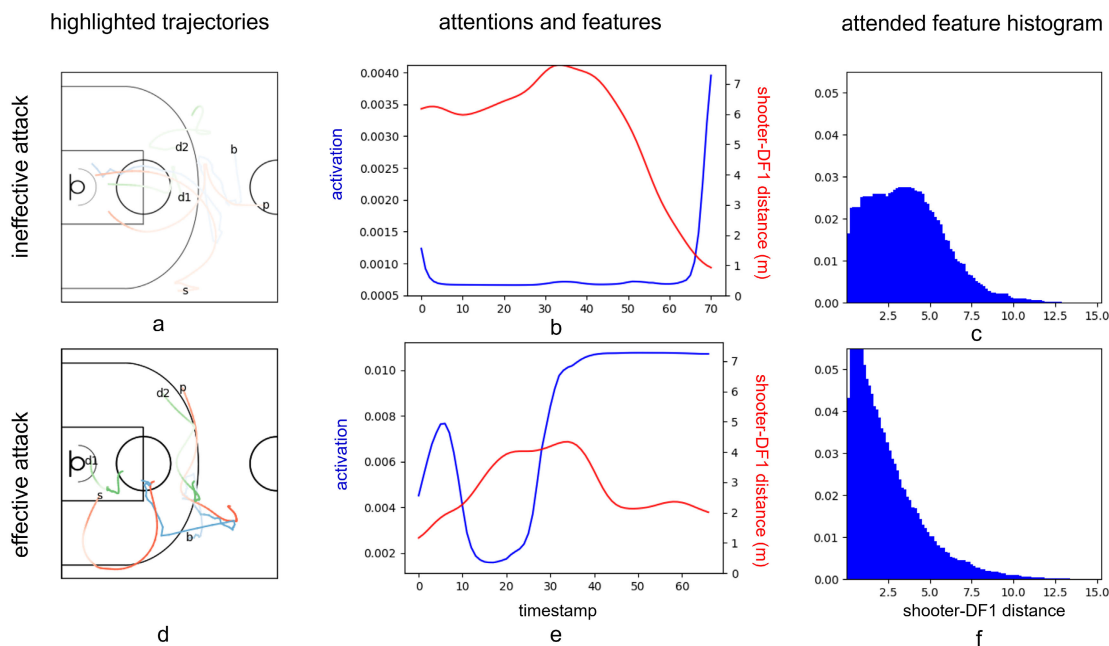
FIGURE 3. Example results of MADCA in effective and ineffective attacks. (a,d) Example highlighted trajectories on a basketball court, (b, e) Example attention values and feature sequences, and (c,f) Distinctive attended feature histogram for a test dataset. In the highlighted trajectories, blue, red, and green represent the ball, attacker, and defender when the features are distinctive between the labels (otherwise, they are white). Attention sequences (blue) are presented with a specific feature (red). The attended feature histogram is based on the distinctive features between the labels during the specific (highlighted) interval.

values within the highlighted segments. Distinctive features were selected based on this value and plotted a histogram to understand the attended (or highlighted) features (Fig. 1d right).

Finally, for team analysis, the average number of effective attacks and goals was analyzed to examine the effective attack label as a team evaluation metric. Pearson correlation coefficients (r-value) was computed between statistical results (e.g., actual goals and effective attacks) and the 2015-16 NBA season results (field goal percentage, and field goal scores), which were obtained from the official NBA website (nba.com). For season results, It was confirmed that field goal percentage was very highly correlated with the season ranking ($\tau = -0.990$ using Kendall's τ), which suggests the field goal percentage reflects the team winning, whereas the field goal scores reflect more offensive aspects. Since the sample size was small ($N = 30$) in the correlation analysis, the r value was used as an effect size for evaluation, rather than the p-value. As described in a previous study [59], correlation coefficients less than 0.20 can be interpreted as *slight, almost negligible relationships*, between 0.20 and 0.40 as *low correlation*; between 0.40 and 0.70 as *moderate correlation*; between 0.70 and 0.90 as *high correlation*, and correlation greater than 0.90 as *very high correlation*.

IV. RESULTS

The purpose of our experiments was to validate our methods for application to real-world team sports data. To this end, our methods were first validated in terms of classification

performance. Then, example visual analyses were presented using our framework. Lastly, quantitative analyses of team performance were shown using our approaches.

A. MODEL VALIDATION

First, our methods were validated in terms of classification performance. Our approach was verified by comparing two baselines with effective/ineffective attack labels and goal/non-goal labels, using different sizes of the dataset (1,024, 8,192, and all 45,307 samples), as shown in Fig. 2. First, as the size of the datasets increased, the prediction performances increased in all models and predictions, indicating that all models would benefit from a greater amount of data. Compared with the goal/non-goal prediction models, effective/ineffective attack prediction models show better performance, which seems reasonable because goal/non-goal prediction is inherently more difficult than effective/ineffective attack prediction. Thus, we basically used the effective/ineffective attack prediction model was basically used with all data for the following analysis. Among the three models, in the effective/ineffective attack prediction, the performance was better in descending order of RNN, CNN-RNN (MADCA-Net), and CNN. In the goal/non-goal prediction, the differences among the models were similar. The prediction performance of the RNN was better than that of MADCA-Net and CNN, but to find the distinctive (highlighted) part of trajectories was aimed, which can be modeled by a 1D CNN. Then, MADCA-Net was used for the

following analysis, which combined the interpretability of the 1D-CNN model and the predictability of an RNN model.

B. EXAMPLE ANALYSIS

Next, example visual analyses were shown using our framework. In this subsection, example results of effective/ineffective attacks were presented in Fig 3, and then those of goal/no-goal attacks were presented in Fig 4.

First, distinctive layers were found by computing a score for each layer by providing a ranking of the layers based on the calculated scores. In both effective/ineffective and goal/no-goal attacks, the last layer of the fourth 1D-CNN is the distinctive layer, which means that the longest filter width in 1D CNN layers was selected.

Next, trajectories colored by the identified distinctive layer were compared. In the example of Fig. 3d (colored by a distinctive layer with the highest score), the start and end segments of an effective attack trajectories were highlighted in color. On the other hand, in the example of Fig. 3a, almost no trajectory segments in an ineffective attack were highlighted. Qualitatively, the latter (Fig. 3a) may be a usual attacking play from the top position (and, therefore, not distinctive), while the former (Fig. 3d) may be an effective shooter movement for creating a scoring opportunity.

Lastly, we tried to understand the meaning of the highlighted segments. MADCA offers two methods for comprehending the rationale behind the attention drawn to a specific segment through a distinctive layer. First, a correlation is computed between the time series of attention values and each of the pre-computed handcrafted features. In effective/ineffective attacks, the distance between the shooter and DF1 (the shooter defender at the last frame) was selected, which seems reasonable because the shooter-DF1 distance is related to shot performance. In Figs 3b and e, activation in attention and the shooter-DF1 distance were negatively correlated, which indicates that MADCA-Net focused on the scoring opportunities with larger shooter-DF1 distances. Second, the difference was provided in distributions of each handcrafted feature among the two classes within the highlighted segments. In Figs 3 c and f, the histograms of the attended feature (in this case, the shooter-DF1 distance) were different between effective and ineffective attacks, which indicates that the attended feature can distinguish between the effective and ineffective attacks.

Next, example MADCA results of goal/no-goal attacks was shown in Fig 4. In the example of Fig. 4a, the start and end segments of an effective attack trajectories were highlighted in color. On the other hand, in the example of Fig. 4a, almost no trajectory segments were highlighted in a no-goal play. Qualitatively, this may not provide useful information because we want to know the highlighted trajectories (i.e., movements) in goal plays rather than no-goal plays. As a distinctive feature, in goal/no-goal attacks, the DF1 y-coordinate was selected, which seems somewhat reasonable but not essential information about multi-agent interaction because this feature indicates that a shooter

TABLE 1. Rankings and statistics of teams in the 2015-2016 NBA season including season field goal (FG) percentages and points, as well as statistical results (mean actual goals and effective attacks from our data).

Team Name	Rank	Season FG %	Season FG pts	Mean goal	Mean eff. att.
Golden State Warriors	1	48.7	8055	0.413	0.478
San Antonio Spurs	2	48.4	7148	0.443	0.512
Oklahoma City Thunder	3	47.6	7422	0.417	0.463
Miami Heat	4	47.0	6798	0.416	0.493
Milwaukee Bucks	5	46.7	6730	0.400	0.501
Los Angeles Clippers	6	46.5	7079	0.395	0.468
Minnesota Timberwolves	7	46.4	6645	0.398	0.466
Sacramento Kings	8	46.4	7226	0.390	0.532
Washington Wizards	9	46.0	7185	0.383	0.505
Cleveland Cavaliers	10	46.0	7222	0.398	0.490
Atlanta Hawks	11	45.8	7151	0.398	0.588
Orlando Magic	12	45.5	7120	0.396	0.522
Brooklyn Nets	13	45.3	6803	0.433	0.466
Houston Rockets	14	45.2	7066	0.387	0.517
Toronto Raptors	15	45.1	6720	0.403	0.483
Portland Trail Blazers	16	45.0	7198	0.403	0.453
Indiana Pacers	17	45.0	6947	0.386	0.460
Utah Jazz	18	44.9	6608	0.389	0.496
New Orleans Pelicans	19	44.8	7008	0.408	0.533
Dallas Mavericks	20	44.4	6934	0.387	0.506
Denver Nuggets	21	44.2	6842	0.399	0.505
Chicago Bulls	22	44.1	6981	0.387	0.460
Memphis Grizzlies	23	44.0	6542	0.386	0.517
Boston Celtics	24	43.9	7149	0.387	0.491
Detroit Pistons	25	43.9	6962	0.369	0.477
New York Knicks	26	43.9	6654	0.422	0.388
Charlotte Hornets	27	43.9	6945	0.414	0.501
Phoenix Suns	28	43.5	6840	0.398	0.505
Philadelphia 76ers	29	43.1	6704	0.364	0.502
Los Angeles Lakers	30	41.4	6399	0.359	0.407

defender retreated toward the end-line of the court but there is no information about the shooter. In fact, there was almost no correlation between activation in attention and the DF1 y-coordinates in Figs 4 b and e. In Figs 4 c and f, the histograms of the attended feature (in this case, DF1's y-coordinate) were almost the same in goal and no-goal plays, which indicates that it would be difficult to distinguish between the two types of plays using the attended feature. Note that, in the above two cases, the information of all features (e.g., passer and DF2) were considered in this analysis. According to the procedure in Section III-E, we selected and showed the distinctive features.

C. TEAM ANALYSIS

Lastly, quantitative analyses of team performance is shown to examine the effective attack metric.

Table 1 shows the 2015-2016 rankings and season field goal percentages and points, as well as statistical results (e.g., actual scores and effective attacks) for each team. The season performance (1,230 games) was estimated using tracking data from a subset of the season's games (600 games).

To gauge the importance of these metrics, using the results from Table 1, Pearson correlation coefficients was computed between the statistical results (actual goals and effective attacks) and 2015-16 NBA season results (field goal (FG) percentage and points). Note that the Golden State Warriors
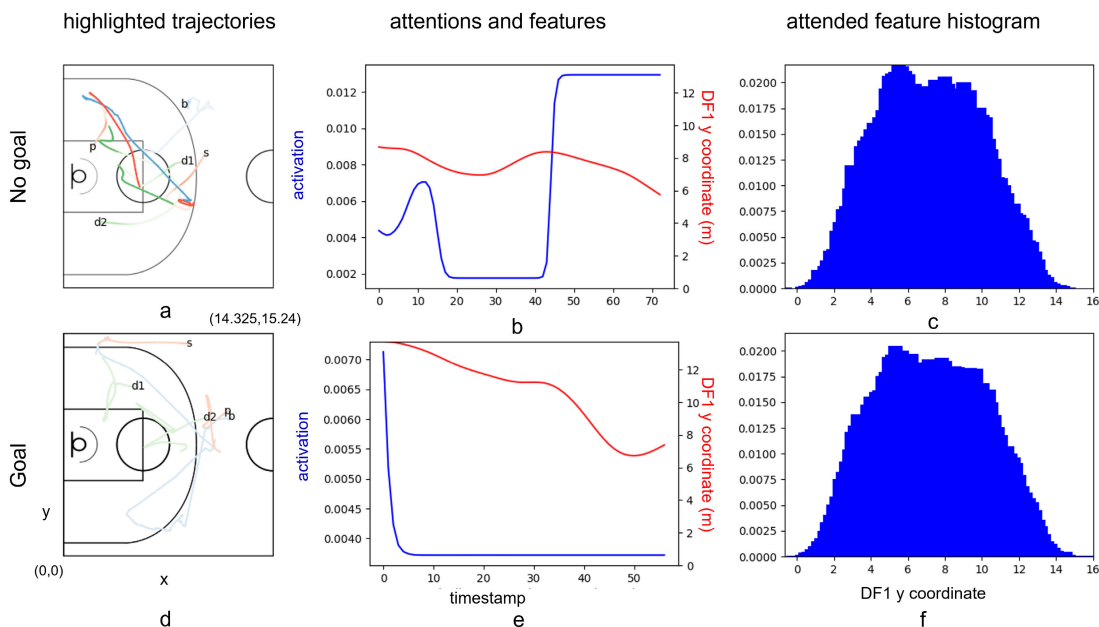
FIGURE 4. Example results of MADCA in goal/no-goal attacks. (a,d) Example highlighted trajectories on a basketball court, (b, e) Example attentions and feature sequences, and (c,f) Distinctive attended feature histograms for a test dataset are shown. Color configurations are the same as those in Fig 1.

TABLE 2. Pearson's *r*-values with each of the quantitative metrics from our data and team performance in the 2015-2016 season (excluding the Warriors).

Quantitative Metric	Season FG (%)	Season FG points
Mean goals	0.595	0.296
Mean effective attacks	0.237	0.349

was excluded, who had a record win-to-loss ratio (73 wins and 9 losses) at the time, from the analysis because the Warriors had, by far, the highest season FG points (8,055) in the league (the second highest was Oklahoma City Thunder with 7,422). From Table 2, the results show that, mean actual goals had moderate and low positive relationships with season FG percentage ($r = 0.595$) and points ($r = 0.296$), respectively. Mean effective attacks, on the other hand, had low positive correlations with season FG percentage ($r = 0.237$) and points ($r = 0.349$), respectively. From these results, the mean FG goal from our data can estimate season FG percentage, which seems reasonable given it is the same data, and the mean effective attacks can estimate season FG points better than the mean FG goals from our data. The results may be related to the effective attack considering the shot area (including 2- and 3-points). Note that, from the correlation results, it is difficult to directly examine the effectiveness of an effective attack because there is no ground truth of an effective attack without considering the scoring results. In other words, the correlation is examined with scoring results but the higher correlation does not mean higher reliability. These results are discussed in the next section.

V. CONCLUSION

In this study, a comparative analysis method called MADCA was proposed, which analyzes multi-agent trajectories in ball games. Our approach was verified by comparing various baselines with effective/ineffective attack labels and goal/non-goal labels, using different sizes of the dataset. The effectiveness of our method was also demonsteated through use cases that analyze the difference between effective and ineffective attacks in the NBA dataset. In this section, our experimental results, the potential limitations of this study, and avenues for future work are discussed.

Our approach can extract distinctive layers in MADCA-Net and features in the effective/ineffective attack prediction task, as shown in Fig 3. However, specifically, for goal/non-goal prediction, when the trajectories of ball sports are dealt with, note that all trajectories may not have the characteristics of a specific class. For example, the shooting skills of a shooter and the randomness of the successful shot affect the goal/non-goal prediction performance. In addition, we can speculate that in top-level teams (e.g., Warriors), players can score even in ineffective situations because of their superior shooting skills. Although it is inherently difficult to validate the effective/ineffective attack metric as mentioned above, we believe it would be more plausible than the metric based on goal/non-goal prediction from these prediction performances. For future work, in trajectories in ball sports, there can be dynamic and detailed labels such as team plays (e.g., screen plays [41], [60]). Utilization of such dynamic and detailed labels may be expected to provide more practical and interpretable results.

In this study, five agents (four players and the ball) were considered to improve interpretability and avoid the role assignment problem. However, in more general cases, all agents (in basketball, 11 agents) should be considered. A graph neural network [49], [51], [61] could be applied to improve predictability and a Gaussian mixture model [47], [50] to address the role assignment problem in order to improve the interpretability of this problem. Combining improvements in predictability and interpretability remains an avenue for future work.

REFERENCES

[1] K. Fujii, N. Takeishi, M. Hojo, Y. Inaba, and Y. Kawahara, "Physically-interpretable classification of biological network dynamics for complex collective motions," *Sci. Rep.*, vol. 10, no. 1, p. 3005, Feb. 2020.

[2] K. Fujii, N. Takeishi, B. Kibushi, M. Kouzaki, and Y. Kawahara, "Data-driven spectral analysis for coordinative structures in periodic human locomotion," *Sci. Rep.*, vol. 9, no. 1, pp. 1–14, Nov. 2019.

[3] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 51, no. 5, p. 4282, 1995.

[4] K. Yokoyama, H. Shima, K. Fujii, N. Tabuchi, and Y. Yamamoto, "Social forces for team coordination in ball possession game," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 97, no. 2, Feb. 2018, Art. no. 022410.

[5] W. Spearman, A. Basye, G. Dick, R. Hotovy, and P. Pop, "Physics-based modeling of pass probabilities in soccer," in *Proc. MIT Sloan Sports Anal. Conf.*, 2017.

[6] F. P. Alguacil, P. P. N. Arce, D. Sumpter, and J. Fernandez, "Seeing in to the future: Using self-propelled particle models to aid player decision-making in soccer," in *Proc. MIT Sloan Sports Anal. Conf.*, 2020.

[7] K. Fujii, K. Yokoyama, T. Koyama, A. Rikukawa, H. Yamada, and Y. Yamamoto, "Resilient help to switch and overlap hierarchical subsystems in a small human group," *Sci. Rep.*, vol. 6, no. 1, pp. 1–10, Apr. 2016.

[8] K. Fujii, T. Kawasaki, Y. Inaba, and Y. Kawahara, "Prediction and classification in equation-free collective motion dynamics," *PLOS Comput. Biol.*, vol. 14, no. 11, Nov. 2018, Art. no. e1006545.

[9] K. Fujii, "Data-driven analysis for understanding team sports behaviors," *J. Robot. Mechtron.*, vol. 33, no. 3, pp. 505–514, Jun. 2021.

[10] T. Maekawa, K. Ohara, Y. Zhang, M. Fukutomi, S. Matsumoto, K. Matsumura, H. Shidara, S. J. Yamazaki, R. Fujisawa, K. Ide, N. Nagaya, K. Yamazaki, S. Koike, T. Miyatake, K. D. Kimura, H. Ogawa, S. Takahashi, and K. Yoda, "Deep learning-assisted comparative analysis of animal trajectories with DeepHL," *Nature Commun.*, vol. 11, no. 1, pp. 1–15, Oct. 2020.

[11] J. Bourbousson, C. Sève, and T. McGarry, "Space–time coordination dynamics in basketball: Part 2. The interaction between the two teams," *J. Sports Sci.*, vol. 28, no. 3, pp. 349–358, Feb. 2010.

[12] B. Travassos, D. Araújo, R. Duarte, and T. McGarry, "Spatiotemporal coordination behaviors in futsal (indoor football) are guided by informational game constraints," *Hum. Movement Sci.*, vol. 31, no. 4, pp. 932–945, Aug. 2012.

[13] J. Sampaio, T. McGarry, J. Calleja-González, S. J. Sáiz, X. S. I. del Alcázar, and M. Balciunas, "Exploring game performance in the national basketball association using player tracking data," *PLoS ONE*, vol. 10, no. 7, Jul. 2015, Art. no. e0132894.

[14] K. Goldsberry, "CourtVision: New visual and spatial analytics for the NBA," in *Proc. MIT Sloan Sports Anal. Conf.*, vol. 9, 2012, pp. 12–15.

[15] V. Correia, D. Araújo, C. Craig, and P. Passos, "Prospective information for pass decisional behavior in rugby union," *Hum. Movement Sci.*, vol. 30, no. 5, pp. 984–997, Oct. 2011.

[16] L. Vilar, D. Araújo, K. Davids, V. Correia, and P. T. Esteves, "Spatial–temporal constraints on decision-making during shooting performance in the team sport of futsal," *J. Sports Sci.*, vol. 31, no. 8, pp. 840–846, Apr. 2013.

[17] K. Fujii, Y. Yoshihara, Y. Matsumoto, K. Tose, H. Takeuchi, M. Isobe, H. Mizuta, D. Maniwa, T. Okamura, T. Murai, Y. Kawahara, and H. Takahashi, "Cognition and interpersonal coordination of patients with schizophrenia who have sports habits," *PLoS ONE*, vol. 15, no. 11, Nov. 2020, Art. no. e0241863.

[18] R. Marty, "High-resolution shot capture reveals systematic biases and an improved method for shooter evalutation," in *Proc. MIT Sloan Sports Anal. Conf.*, 2018.

[19] A. Miller, L. Bornn, R. Adams, and K. Goldsberry, "Factorized point process intensities: A spatial analysis of professional basketball," in *Proc. Int. Conf. Mach. Learn.*, 2014, pp. 235–243.

[20] E. Papalexakis and K. Pelechrinis, "THoops: A multi-aspect analytical framework for spatio-temporal basketball data," in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2018, pp. 2223–2232.

[21] Q. Wang, H. Zhu, W. Hu, Z. Shen, and Y. Yao, "Discerning tactical patterns for professional soccer teams: An enhanced topic model with applications," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2015, pp. 2197–2206.

[22] A. C. Miller and L. Bornn, "Possession sketches: Mapping NBA strategies," in *Proc. MIT Sloan Sports Anal. Conf.*, 2017.

[23] K.-C. Wang and R. Zemel, "Classifying NBA offensive plays using neural networks," in *Proc. MIT Sloan Sports Anal. Conf.*, 2016.

[24] A. Nistala, "Using deep learning to understand patterns of player movement in basketball," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 2018.

[25] A. Grunz, D. Memmert, and J. Perl, "Tactical pattern recognition in soccer games by means of special self-organizing maps," *Hum. Movement Sci.*, vol. 31, no. 2, pp. 334–343, Apr. 2012.

[26] M. Kempe, A. Grunz, and D. Memmert, "Detecting tactical patterns in basketball: Comparison of merge self-organising maps and dynamic controlled neural networks," *Eur. J. Sport Sci.*, vol. 15, no. 4, pp. 249–255, May 2015.

[27] K. Fujii, Y. Inaba, and Y. Kawahara, "Koopman spectral kernels for comparing complex dynamics: Application to multiagent sport plays," in *Proc. Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Cham, Switzerland: Springer, 2017, pp. 127–139.

[28] J. Hobbs, P. Power, and L. Sha, "Quantifying the value of transitions in soccer via spatiotemporal trajectory clustering," in *Proc. MIT Sloan Sports Anal. Conf.*, 2018.

[29] J. Hobbs, M. Holbrook, N. Frank, L. Sha, and P. Lucey, "Improved structural discovery and representation learning of multi-agent data," 2019, *arXiv:1912.13107*.

[30] T. Decroos, J. Van Haaren, and J. Davis, "Automatic discovery of tactics in spatio-temporal soccer match data," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 223–232.

[31] L. Sha, P. Lucey, Y. Yue, P. Carr, C. Rohlf, and I. Matthews, "Chalkboarding: A new spatiotemporal query paradigm for sports play retrieval," in *Proc. 21st Int. Conf. Intell. User Interfaces*, Mar. 2016, pp. 336–347.

[32] S. Kanda, K. Takeuchi, K. Fujii, and Y. Tabei, "Succinct trit-array trie for scalable trajectory similarity search," in *Proc. 28th Int. Conf. Adv. Geograph. Inf. Syst.*, Nov. 2020, pp. 518–529.

[33] K. Takeuchi, M. Imaizumi, S. Kanda, Y. Tabei, K. Fujii, K. Yoda, M. Ishihata, and T. Maekawa, "Fréchet kernel for trajectory data analysis," in *Proc. 29th Int. Conf. Adv. Geograph. Inf. Syst.*, 2021, pp. 221–224.

[34] M. A. Alcorn and A. Nguyen, "baller2vec: A multi-entity transformer for multi-agent spatiotemporal modeling," 2021, *arXiv:2102.03291*.

[35] D. Cervone, A. D'Amour, L. Bornn, and K. Goldsberry, "PointWise: Predicting points and valuing decisions in real time with NBA optical tracking data," in *Proc. MIT Sloan Sports Anal. Conf.*, 2014.

[36] Y.-H. Chang, R. Maheswaran, S. J. J. Kwok, T. Levy, A. Wexler, and K. Squire, "Quantifying shot quality in the NBA," in *Proc. MIT Sloan Sports Anal. Conf.*, 2014.

[37] P. Lucey, A. Bialkowski, M. Monfort, P. Carr, and I. Matthews, "Quality vs quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data," in *Proc. MIT Sloan Sports Anal. Conf.*, 2014.

[38] P. Lucey, D. Oliver, P. Carr, J. Roth, and I. Matthews, "Assessing team strategy using spatiotemporal data," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2013, pp. 1366–1374.

[39] A. McQueen, J. Wiens, and J. Guttag, "Automatically recognizing on-ball screens," in *Proc. MIT Sloan Sports Anal. Conf.*, 2014.

[40] A. McIntyre, J. Brooks, J. Guttag, and J. Wiens, "Recognizing and analyzing ball screen defense in the NBA," in *Proc. MIT Sloan Sports Anal. Conf.*, 2016, pp. 11–12.

[41] M. Hojo, K. Fujii, Y. Inaba, Y. Motoyasu, and Y. Kawahara, "Automatically recognizing strategic cooperative behaviors in various situations of a team sport," *PLoS ONE*, vol. 13, no. 12, Dec. 2018, Art. no. e0209247.

[42] M. Hojo, K. Fujii, and Y. Kawahara, "Analysis of factors predicting who obtains a ball in basketball rebounding situations," *Int. J. Perform. Anal. Sport*, vol. 10, pp. 1–14, Jan. 2019.

[43] R. Bunker, K. Fujii, H. Hanada, and I. Takeuchi, "Supervised sequential pattern mining of event sequences in sport to identify important patterns of play: An application to rugby union," *PLoS ONE*, vol. 16, no. 9, Sep. 2021, Art. no. e0256329.

[44] N. Mehrasa, Y. Zhong, F. Tung, L. Bornn, and G. Mori, "Deep learning of player trajectory representations for team activity analysis," in *Proc. MIT Sloan Sports Anal. Conf.*, 2018.

[45] A. Sicilia, K. Pelechrinis, and K. Goldsberry, "DeepHoops: Evaluating micro-actions in basketball using deep feature representations of spatio-temporal data," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 2096–2104.

[46] S. Zheng, Y. Yue, and J. Hobbs, "Generating long-term trajectories using deep hierarchical networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 1543–1551.

[47] H. M. Le, Y. Yue, P. Carr, and P. Lucey, "Coordinated multi-agent imitation learning," in *Proc. 34th Int. Conf. Mach. Learn. 70*, 2017, pp. 1995–2003.

[48] M. Teranishi, K. Fujii, and K. Takeda, "Trajectory prediction with imitation learning reflecting defensive evaluation in team sports," in *Proc. IEEE 9th Global Conf. Consum. Electron. (GCCE)*, Oct. 2020, pp. 124–125.

[49] M. Teranishi, K. Tsutsui, K. Takeda, and K. Fujii, "Evaluation of creating scoring opportunities for teammates in soccer via trajectory prediction," in *Proc. 9th Workshop Mach. Learn. Data Mining Sports Anal. Eur. Conf. Mach. Learn. Princ. Pract. Knowl. Discovery (ECML-PKDD)*, 2022, pp. 53–73.

[50] K. Fujii, N. Takeishi, Y. Kawahara, and K. Takeda, "Policy learning with partial observation and mechanical constraints for multi-person modeling," 2020, *arXiv:2007.03155*.

[51] K. Fujii, K. Takeuchi, A. Kuribayashi, N. Takeishi, Y. Kawahara, and K. Takeda, "Estimating counterfactual treatment outcomes over time in complex multi-agent scenarios," 2022, *arXiv:2206.01900*.

[52] Y. Yue, P. Lucey, P. Carr, A. Bialkowski, and I. Matthews, "Learning fine-grained spatial models for dynamic sports play prediction," in *Proc. IEEE Int. Conf. Data Mining*, Dec. 2014, pp. 670–679.

[53] S. Zheng, R. Yu, and Y. Yue, "Multi-resolution tensor learning for large-scale spatial data," 2018, *arXiv:1802.06825*.

[54] J. Y. Park, K. T. Carr, S. Zheng, Y. Yue, and R. Yu, "Multiresolution tensor learning for efficient and interpretable spatial analysis," 2020, *arXiv:2002.05578*.

[55] J. Mortensen and L. Bornn, "From Markov models to Poisson point processes: Modeling movement in the NBA," in *Proc. MIT Sloan Sports Anal. Conf.*, 2019.

[56] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder–decoder approaches," 2014, *arXiv:1409.1259*.

[57] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.

[58] (2015). *NBA Advanced Stats*. Accessed: Mar. 18, 2023. [Online]. Available: https://www.nba.com/stats/players/shots-closest-defender/

[59] J. P. Guilford, *Fundamental Statistics in Psychology and Education*. New York, NY, USA: McGraw-Hill, 1950.

[60] Z. Ziyi, K. Takeda, and K. Fujii, "Cooperative play classification in team sports via semi-supervised learning," *Int. J. Comput. Sci. Sport*, vol. 21, no. 1, pp. 111–121, Mar. 2022.

[61] R. A. Yeh, A. G. Schwing, J. Huang, and K. Murphy, "Diverse generation for multi-agent sports games," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4610–4619.

ZHANG ZIYI received the B.E. degree in computer science and technology from Southeast University, Nanjing, China, in 2017, and the M.E. degree in production systems from Waseda University, Kitakyushu, Japan, in 2018. He is currently pursuing the Ph.D. degree in intelligent systems with Nagoya University, Nagoya, Japan. His research interest includes team sport trajectory analysis.

RORY BUNKER received the B.Com. and B.Com. (Hons.) degrees from The University of Auckland, New Zealand, in 2010 and 2011, respectively, and the master's degree in analytics from the Auckland University of Technology, New Zealand, in 2016. He is currently a member of the Sports Behavior Group, Department of Intelligent Systems, Graduate School of Informatics, Nagoya University, Japan. His current research interests include trajectory and pattern mining in sports and machine learning for sports performance analysis and outcome prediction.

KAZUYA TAKEDA (Senior Member, IEEE) received the B.E., M.E., and Ph.D. degrees from Nagoya University. He was with the Advanced Telecommunication Research Laboratories and the KDD R&D Laboratories. In 1995, he joined Nagoya University, where he started a research group for signal processing applications. He is currently a Professor with the Graduate School of Informatics and the Green Mobility Collaborative Research Center, Nagoya University. His research interests include behavior signal processing and driving behavior.

KEISUKE FUJII (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Kyoto University, in 2009, 2011, and 2014, respectively. After his work, he was a Postdoctoral Fellow and a Research Scientist with Nagoya University and the RIKEN Center for Advanced Intelligence Project, Japan, he joined Nagoya University. He is currently an Associate Professor with the Graduate School of Informatics. His research interests include interdisciplinary studies among machine learning, behavioral sciences, and sports sciences.

• • •