AI-Driven Assessment of Player Contribution and Cooperation Dynamics in Team Sports Using Trajectory Data

Submitted to
Graduate School of Information Science and Technology
Osaka University

January 2025

Takeshi TANAKA

List of Publications

Related Articles

(1) Takeshi Tanaka, Takuya Magome, Shunichi Tazuke, Norio Goda, "Analysis of Potential Information Transmission with Transfer Entropy in Team Sports," *IPSJ Journal*, vol. 62, no. 2, pp. 737-746, February 2021. (in Japanese)

(2) Takeshi Tanaka, Touru Yamaguchi, Shunichi Tazuke, Norio Goda, "Feature Extraction Method from Motion Data for Risk Assessment of Sports Injuries," *IPSJ Journal*, vol. 63, no. 7, pp. 1321-1330, July 2022. (in Japanese)

Related Conference Papers

(1) Takeshi Tanaka, Akira Uchiyama, Hirozumi Yamaguchi, "ML-based Individual Contribution Assessment of Basketball Players from Their Trajectories," in *Proceedings of the 24th IEEE International Conference on Mobile Data Management (MDM 2023)*, pp. 55-64, August, 2023.

Abstract

The integration of advanced image processing technology and GPS-enabled devices has greatly improved our ability to analyze human movement, and has led to a deeper understanding of group dynamics in various fields, particularly in sports. However, extracting the rationale for the predictions is a challenge in AI research that aims to achieve advanced predictions from high-dimensional information. In this paper, we examine an advanced approach to predicting the performance of a sports team and the contribution of individual players By applying a method of adding perturbations to the input of a black box AI model. Our research aims to provide a highly interpretable prediction process by using a deep learning model that takes as input the trajectories of players to evaluate both the performance of individual players and the interactions between players in team sports such as basketball and football.

We introduce a methodology that consists of three comprehensive approaches. First, we develop a convolutional neural network (CNN) model that estimates the contribution of players using perturbations of the input data. This model predicts key performance indicators such as team scores by analyzing the trajectories of players and the ball. Furthermore, this model provides a fundamental tool for understanding how individual actions contribute to the success of the team as a whole. Next, we delved into evaluating the interactions between players by applying the principles of information theory, specifically using transfer entropy to quantify the degree of cooperation between players based on their movement patterns. This analysis provides insight into how the interactions between players affect team dynamics and outcomes. Third, we enhanced the predictive model by incorporating these interaction insights using a graph neural network (GNN). This allows us to more accurately reproduce the relational dynamics between all players and the ball, ultimately improving prediction accuracy.

We rigorously verified the method we proposed using the vast amounts of tracking data from professional basketball and football leagues. This resulted in a high-precision scoring prediction model for basketball, achieving an area under the curve (AUC) of approximately 0.92. The contribution of each player was evaluated by assessing the change in output values when the input of each player was masked, and the validity of the results was confirmed by comparing the average values for one season with existing evaluation metrics. In particular, it was found that the correlation with the efficiency, a measure of the contribution of basketball players, was significantly high, while the correlation with other individual scores such as points and assists was low. A model was also constructed for predicting

scoring opportunities in soccer, and a prediction accuracy of approximately 0.87 was achieved, although the dataset was smaller than that for basketball. In addition, when evaluating the contribution of players to scoring opportunities in soccer, we evaluated the correlation with individual scores for each team, and found that it was highly correlated with the number of offensive plays that had a large impact on the movement trajectory, such as shooting, receiving passes, and dribbling. From these results, we found that the proposed player contribution evaluation is a new method that evaluates the direct contribution to scoring using explainable AI, and that it is also possible to interpret the validity from the conventionally used indicators.

In the second method, we used the concept of transfer entropy from information theory to calculate the degree of cooperation between players in a soccer match. To calculate this, we proposed a method of converting time series data such as the independent acceleration and movement of two people into continuous probability variables, and evaluating whether there is cooperation between the two people. In addition, in order to collect evaluation data in actual team sports, we designed and prototyped a wearable device that can collect multiple sensor data with little burden on the players. Using the data collected from the football matches using the prototype device, we compared the network structure based on transfer entropy with the organised performance of the matches as evaluated by experts. As a result, we were able to confirm that cooperation between teams in particular has a significant impact on organised performance.

In the third method, we applied a Graph Neural Network (GNN) to the deep learning model of the first method based on the hypothesis identified in the second method. By incorporating interactions between players, we aimed to improve the accuracy of the basketball scoring prediction model and improve the validity of the player contribution evaluation with a more realistic model. Using a basketball dataset, we evaluated the prediction accuracy using three different assumptions about the graph structure between players. As a result, we confirmed that the highest-precision scoring prediction model, with an AUC of 0.95 or higher, was obtained by considering the interaction between all players and the ball. In addition, we confirmed that the prediction accuracy was higher than that of the model using CNN that did not consider the interaction between players, and that the precision and recall were also close to the level required in sports, at around 0.9.

The research results highlight the potential for AI to automate and enhance traditional evaluation methods used by coaches and sports analysts. By providing a comprehensive framework for evaluating player contributions and interactions based on data, it can facilitate the development of innovative tactics and enable the re-evaluation of previously overlooked players. This will have practical implications for team strategy and player development, providing a data-driven basis for decision-making. Beyond immediate application in the field of sport, this research has the potential for wider impact, including in the training of athletes, the assessment of performance and the understanding of developmental processes. These outcomes will make a significant contribution to educational and lifelong sporting initiatives, highlighting the potential for data analysis techniques to positively impact on children's education and overall health.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The ability to quantitatively analyze group movements using data gathered from cameras and smart devices has become a cornerstone of modern technological advancements. This progress is largely due to developments in image analysis technology, which now allow for the precise recognition of people and objects in two-dimensional images and their conversion into location data. Additionally, the miniaturization of devices equipped with GPS and communication capabilities, such as smartphones and smartwatches, has facilitated the routine collection of data on individuals' locations and movements [1, 2]. Researchers are actively investigating methods for analyzing this location and movement data using various algorithms and AI techniques, including machine learning [3]. Insights derived from understanding human behavior in urban environments and buildings are being leveraged to design more efficient transportation systems and better structures [4, 5]. Moreover, analyzing employee behavior in corporate settings is being utilized to boost productivity. In educational contexts and team sports, such analyses not only aim to improve gymnastics education and strengthen competitive teams but also serve as model cases for research into group behavior by defining specific rules and environments [6].

In team sports research, team performance largely depends on individual players' performances and their interactions. Historically, studies have focused on two main areas: physical and technical performance. Initially, wearable sensors equipped with GPS and accelerometers were introduced to professional teams like football, allowing for the evaluation of physical performance metrics such as the distance covered and speed during games and training sessions [7]. Long-term observation and analysis of these physical indicators have led to the development of training methods that maximize physical performance. Research on technical skills began with statistical evaluations of shots and dribbles in sports like basketball and football [8]. More recent studies have focused on skills that significantly impact game outcomes, such as shot difficulty, estimated from a player's position and the ball's trajectory [9]. Interaction research involves analyzing player positions in football to determine advantageous positions for team play [10]. Various models are studied to estimate how interactions, such as passing exchanges, affect team performance [10]. With advancements in data complexity and

AI technologies like deep learning, it is now possible to predict team performance using primary data such as player behavior and movement. In the case of intrusive sports such as football and basketball, excluding data that expresses the characteristics of the players and the environment, play information is expressed as spatio-temporal data for the players and the ball [10]. Spatio-temporal data can be divided into two types: trajectory data and event data. Trajectory data is a time series of positional coordinates for the players and the ball. Event data is data that records the type of play, such as a player shooting or passing, and the time and location at which it was executed. These describe different aspects of play, and are complementary in that combining them makes it possible to describe the game in more detail. In this study, we focus on trajectory data first, as the main purpose is to provide tactical feedback in team sports. In addition, by including trajectory data for the ball, we can include information equivalent to important event information such as passes and shots, and we have sufficient information for tactical analysis.

In team sports such as basketball and football, tracking data has been used in the past to prevent injuries and plan training, but AI-based predictions using data on the trajectories of players and balls are expected to strengthen teams more tactically [11, 12]. One is to discover and understand plays that contribute to high performance, such as scoring, in order to plan team tactics. By analyzing these AI predictions together with the basis and context behind them, as well as the corresponding video footage, team coaches can use them to develop new tactics that are more suited to their team. The other use is for objective player evaluation. Until now, player evaluation has either involved using statistical indicators such as scoring, or it has required watching huge amounts of video footage. AI-based player evaluation can be used for fair comparisons of player performance within a team, as well as for efficiently discovering and acquiring new promising players. However, previous AI research in team sports has often resulted in models that are considered black boxes, making it challenging to derive insights from the prediction process [13]. These models are typically built upon empirical hypotheses suggesting that team performance consists of individual and collective performances. However, the opaqueness of these models makes it difficult to understand how these factors lead to performance outcomes. Recent efforts have aimed to visualize the analytical evidence within AI models that use high-dimensional data and images, thereby enhancing the reliability of AI predictions [14]. In sports, interpreting AI model estimation processes could yield insights beyond human experiential knowledge.

In this study, we introduce a method for evaluating player performance and interactions between players using an AI model that incorporates collaboration between players based on positional and motion data. Ultimately, we will apply a method that involves adding perturbations to the input data of the AI model and observing changes in the prediction results, in order to predict team performance and interpret the basis for the model's prediction results. First, we will propose a method for estimating the contribution of each player to the basis for the prediction, using a deep learning model that predicts the score, an important indicator of team performance, from the trajectories of players and the ball in sports such as basketball and soccer. Next, in order to construct a model that can more realistically estimate team and individual performance, we evaluate the interactions between players based on information theory, and in particular, we examine the causal relationships between the movements of two players in detail. Finally, we extend the first method by applying a neural network structure that can learn the mutual positional relationships between players based on the findings of the interactions between players.

To validate these methods, we utilized tracking data from professional basketball and soccer league games. For the first method, we developed a deep learning model using a convolutional neural network (CNN) to estimate scores from player and ball trajectories. This model achieved high precision in basketball scoring predictions, with an area under the curve (AUC) of 0.9 or higher. We evaluated player contributions by observing changes in output values when input data for each player was masked, confirming validity through comparisons with existing player evaluation metrics. A similar model was developed for estimating scoring opportunities in soccer, yielding comparable results. In the second method, we employed 'transfer entropy,' an information theory concept, to calculate player cooperation levels in soccer. By comparing network structures based on transfer entropy and match performance, we confirmed that calculated cooperation levels significantly impact performance. In the third method, building on the second method's hypothesis, we applied a Graph Neural Network (GNN) to the first method's deep learning model, evaluating basketball scoring accuracy by considering player interactions. This evaluation confirmed that considering all players' interactions and the ball leads to a more accurate scoring estimation model.

The findings of this research enable the use of AI models to evaluate each player's direct contribution to team performance, such as scoring. Additionally, by integrating player interactions into the model, team performance predictions become more accurate, resulting in a more precise evaluation of individual contributions. Utilizing such AI models can automate empirical evaluations traditionally conducted by coaches and experts, leading to the development of new tactics and the reevaluation of previously overlooked players. However, this research is limited by the input data and model construction process. Only player trajectory data was used, showing that sports performance can be extracted from positional information, but excluding physical abilities and ball skills not reflected in positional data. Future developments should aim to create AI models that incorporate these additional inputs, ultimately leading to accurate predictions of player behavior in complex team sports. Furthermore, understanding sports performance and development processes can contribute to children's education and lifelong wellbeing.

Chapter 2

Related Work

2.1 Trajectory Mining

Pre-processing and characterization are used to remove information unnecessary for analysis from trajectory data and convert it into a data shape that is easy to compare [3, 15]. Various unsupervised and supervised machine learning methods are used for classification and discrimination using the characterized trajectory data as input [16]. In recent years, research on trajectory mining has mainly focused on the use of deep learning and other machine learning methods to make predictions based on trajectory data and related data, and this data is mainly obtained using smartphone GPS and image recognition. Many studies have been reported on predicting movement routes [17], destinations [18], traffic congestion [19], and movement speed [20], as well as the occurrence of events [21] and human decision-making [22]. These studies are concerned with a single person or a moving object's trajectory, not with predictions using multiple trajectories as input. They do not take into account their mutual impact on each other. In addition, studies dealing with the interaction of trajectories of multiple people have reported studies that discover groups that act together in crowd trajectories [23]. Representative studies aim at predicting movement trajectories and objectives for crowd movement, etc., [24], mining similar groups, and extracting outliers [25]. In an analytical method that uses a graph structure to account for the interaction of multiple people, studies have also been reported that combine GNN and LSTM to predict the trajectory ahead, using the trajectories of multiple pedestrians passing by as input [26]. Similar GCN-based methods have been reported for multiple vehicle trajectory prediction [27]. However, there was no focus on the graph structure within the group or even the interpretation of individual contributions.

2.2 Sports Analysis Using Tracking Data

In sports applications, several studies using group trajectories to use trajectory information for tactics and training have been reported, as well as studies on visualization and retrieval methods for trajectory information [28, 29]. In addition, studies have been reported on event detection, prediction of movement direction and play area, and prediction of game results in team sports using machine learning with trajectory data as input [30, 31]. Studies have also been reported on classifying team formations, estimating roles [32, 33], and identifying group plays and tactics [34, 35]. These studies help to predict and understand team performance but do not quantify the contribution of individuals within a team. Studies aimed at estimating individual performance have been reported, such as predicting the performance of athletes in cycling, an individual sport [36]. To understand the degree of influence among individuals in team sports, studies have also been reported using transfer entropy to estimate the causal effects of movement among multiple athletes and using network structure to understand group performance [37]. Although hypotheses have been proposed to interpret the role of individuals in a team, as in this study, they have not been able to directly represent their contribution to team performance.

2.3 Method for Obtaining the Explanatory Basis of Machine Learning Models

Estimation models using deep learning can handle high-dimensional data without manual characterization and provide highly accurate estimation, but the problem is that the estimation process is a black box. Therefore, "explainability," which clarifies the basis of estimation, has become a new research topic called XAI (explainable artificial intelligence) [14]. LIME [38], which evaluates the contribution of input data using a linear model that approximates part of a black-box model, is a representative method. SHAP [39], which uses Sharpe Ray values from cooperative game theory to evaluate contributions, is a similar method. In addition, for deep learning models that use images as input, studies have been reported that visualize and quantify the part of the input image to focus on (often called the point of interest) [40]. Ablation, a method of adding perturbations to the input data, is widely used in the approach of these studies. This method evaluates the change in results when parts of the model are excluded [41].

Ablation-CAM [42], one of the CAM (class activation maps) methods that visualize the points of attention of a model that inputs images such as CNN, is a method that visualizes the points that the model refers to by evaluating the changes caused by excluding the output of some feature maps. These XAI methods are expected to be applied to the use of deep learning models in the medical field, where experts check the evidence in addition to the estimation of deep learning models when making important decisions such as diagnosis. It is also being applied in the development stage of deep learning models and for debugging purposes to improve robustness and search for more accurate models in the practical stage.

CAM is a representative method of XAI in deep learning of image recognition, which highlights input regions that are considered to have contributed to a particular class [43]. While the original

CAM could only be applied to specific models, grad-CAM [44] can be applied to models using CNNs in general-purpose hands, and various improved methods have been reported [45, 46], ablation-cam being one of them. The ablation study approach has also been used in fields other than image recognition. Its application to model understanding in deep learning with time-series data input, such as electroencephalography, has been reported [47]. Similarly, studies have been reported to evaluate the importance of modality in CNN models with time-series multimodal data input [48].

A method of interpreting the basis for predictions is also being researched, whereby the input data is only perturbed and the prediction results are observed, without making any changes to the deep learning model itself, as in the case of CAM. In research on deep learning models that take images as input, a method of adding perturbations by masking parts of the input image has been reported [49, 50, 51]. The idea for our research was inspired by the idea that, as with previous research, it is possible to understand the contribution of the input values by looking at the change in the output values when a perturbation that removes part of the input values is applied.

Chapter 3

ML-based Individual Contribution Assessment of Players in Team Sports from Their Trajectories

3.1 Introduction

The analysis of location information acquired in everyday life enables a detailed understanding of human behavior and contributes significantly to society. In recent years, the development of location information acquisition technologies, such as GPS installed in smartphones and other devices and human recognition using surveillance camera images [1, 2], has made it possible to acquire large amounts of trajectory data, which is continuous location information of people, vehicles, and other objects. The large amount of trajectory data has attracted attention from many fields, including computer science, sociology, and geography, and analysis methods have been actively researched. Understanding behavior by analyzing trajectory data has been applied to transportation systems [4, 5], marketing [52], and health care [53], among others.

Various data processing methods have been developed and applied to understand trajectory data. Preprocessing and characterization are used to remove unnecessary information from trajectory data for analysis and to transform it into a data form that can be easily compared [3, 15]. Various unsupervised and supervised machine learning methods are used for classification and discrimination using the characterized trajectory data as input [16]. In addition, deep learning is used as a method that does not require characterization based on prior human knowledge to reduce analysis person-hours. Deep learning methods for analyzing trajectory data include RNN (Recurrent Neural Network) [54], which learns time series information; CNN (Convolutional Neural Network) [55], which inputs trajectory image data for objects moving in free space such as ships; and methods that combine RNN and CNN [56].

To understand behavior by analyzing trajectory data, recent research has focused on extracting meaning from multiple trajectories based not only on the behavior of individuals but also on data from crowds of unspecified people [57]. Representative studies aim at predicting motion trajectories and goals for crowd motion, etc. [24], mining similar groups, and extracting outliers [25]. In addition, team sports, where multiple people act toward a single goal, are the optimal target for research on understanding group behavior [6]. A spatio-temporal analysis method for group trajectory data has been proposed for scene evaluation in team sports [10]. Methods using deep learning have been investigated, such as score prediction [13], classification of player roles and team styles [58, 59], and simulation of specific scenes with GANs [60], using deep learning of RNNs and CNNs.

However, while previous studies have been able to predict performance, such as scoring chances, from group trajectories, they have not quantitatively assessed the degree of individual influence on group performance. Traditionally, various player evaluation metrics have been developed and combined through empirical knowledge, using scores such as goals and assists, as well as player and ball position information. However, there is no way to predict the direct contribution to scoring or team wins. Over the years, sports data analytics specialists have created and updated new hypotheses and metrics through trial and error.

In this study, we propose a deep learning model that predicts scoring opportunities based on trajectory data for all players and the ball in team sports, as well as a new method for quantitatively evaluating the contribution of individual players based on perturbations to the input of the trained prediction model. In order to implement the proposed method, we design a new neural network structure that makes it easy to mask the input data for each player, unlike previous research. In this method, we generate a dataset for each scene using tracking data for players and the ball, as well as event data such as goals and pass interceptions. Then, players' contribution to scoring opportunities in attacking scenes is quantitatively evaluated. First, a CNN-based deep learning model is trained to predict the probability of scoring opportunities from all player and ball trajectory images. Next, the deep learning model is trained and used to quantify the players' contribution based on the change in the model output values when the player trajectory images are removed from the model input.

In this work, we further generalize our algorithm to handle other team sports with different rules and systems. Specifically, depending on the target sport, we developed a process that optimizes the number of data inputs to the deep learning model and the position of data inputs corresponding to team composition and player roles. Furthermore, in preprocessing input image data, we made it possible to generate trajectory images according to different field settings and outcomes and scene delimitation timing according to event data. To demonstrate its effectiveness, we carried out a new experiment focusing on soccer matches. We utilized the official tracking data from the J1 League, a premier professional soccer league in Japan, to quantify players' contributions in scenarios that resulted in scoring.

We confirmed high predictive performance for both the basketball and soccer models and validated the results by comparing them with existing indices of player contribution. In basketball, we confirmed that prediction accuracy of AUC=0.92 was obtained by dividing the dataset into 100,000 attack scenes extracted from approximately 600 games. When the contribution of each player calculated from the learning model was compared with the official overall player evaluation index, the average annual "efficiency," a significant correlation was confirmed with Spearman's correlation coefficient R=0.37 (p<0.001). In soccer, we confirmed that prediction accuracy of AUC=0.87 was obtained by dividing the dataset into 4,000 attacking scenes extracted from a smaller number of data, approximately 50 matches. In addition, it was confirmed that each player's contribution calculated from the learning model is significantly correlated with indicators of offensive performance, such as the number of shots and pass receptions per year. The evaluation results suggest that the proposed method is valid compared to existing player performance indices and can be a new method to evaluate players based only on trajectory data, in contrast to conventional approaches that rely on multiple numerical parameters that require much effort in data collection and analysis.

3.2 Method

3.2.1 Method Overview

We propose a method for estimating the degree of individual contribution to team performance using a machine learning model that predicts team performance, such as team scores. Our main module is a team performance predictor based on CNNs taking trajectories of players and a ball as input. We assume the trajectories are generated by techniques such as tracking on video data. The team performance predictor predicts the likelihood of scoring using trajectories in the attacking scene as input. Figure 3.1 illustrates the overview of the proposed method. We let $\boldsymbol{A}_i$ denote a trajectory of a player i. To make the processing by CNN easier, each trajectory is represented as a grayscale image. Then, we define a sequence of n trajectories of n players recorded during the same period as:

$$\boldsymbol{A}^n = (\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_n) \tag{3.1}$$

Given $\boldsymbol{A}^n$, the team performance predictor outputs $P(\boldsymbol{A}^n)$. Inspired by the ablation study approach [38,47], our method estimates a loss of the output value due to individual contribution $R(i, \boldsymbol{A}^n)$ of player i for the trajectory sequence $\boldsymbol{A}^n$ by excluding trajectory $\boldsymbol{A}_i$ from the input to the team performance predictor. For this purpose, we multiply the zero matrix $\boldsymbol{O}$ to $\boldsymbol{A}_i$ and replace $\boldsymbol{A}_i$ in $\boldsymbol{A}^n$ with $\boldsymbol{O}\boldsymbol{A}_i$. To be more specific, we define a trajectory sequence $E(i, \boldsymbol{A}^n)$ as:

$$E(i, \boldsymbol{A}^n) = (\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{O}\boldsymbol{A}_i, \cdots, \boldsymbol{A}_n) \tag{3.2}$$

Then, the loss of the output value due to individual contribution $R(i, \boldsymbol{A}^n)$ is defined as below.

$$R(i, \boldsymbol{A}^n) = P(E(i, \boldsymbol{A}^n)) - P(\boldsymbol{A}^n) \tag{3.3}$$

Intuitively, the degree of individual contribution represents the decrease in the loss of output value of the team performance predictor without the player. For estimating the degree of the individual

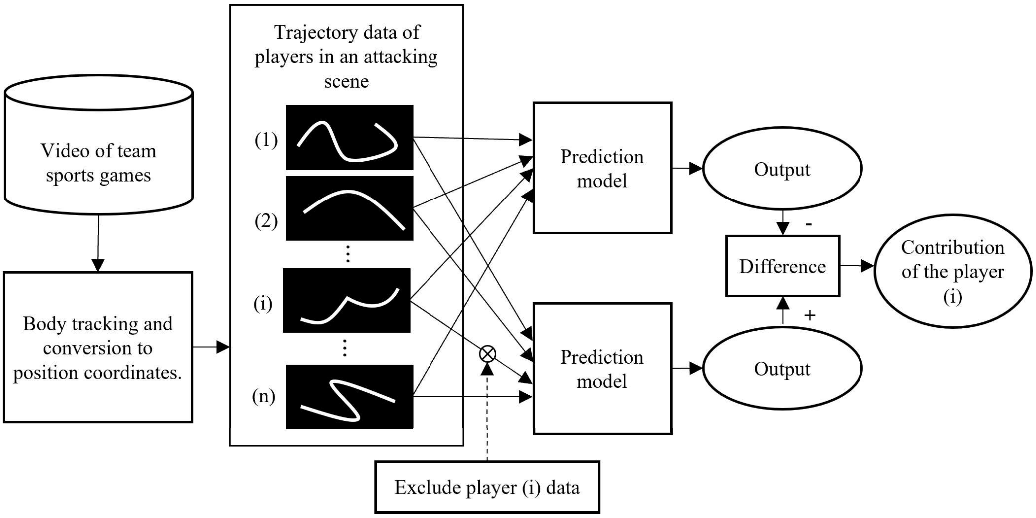


Figure 3.1: Our proposed method for estimating the contribution of individuals in team sports using machine learning models with multiple players' trajectory data as input

contribution of player i, we use trajectory sequences in attacking scenes of i that are predicted as success to score. Note that we also use trajectory sequences that are predicted to score but are not actually scored. We focus only on attacking scenes predicted to score because player movements other than scenes leading to score do not represent good performance. Although not evaluated in this paper, conversely, only scenes where the offensive team failed to score can be used when assessing the contribution to defense. In such a case, the individual contribution can be defined as the increase in the output value when excluding the target player from the defending team.

3.2.2 Neural Network for Score Prediction Model

Figure 3.2 shows our design of the team performance predictor, which predicts success or failure in scoring for the input of a 2D image representing the trajectories of all players and a ball. The network structure of the team performance predictor is based on a previous study [18]. In the post-learning prediction phase, the image is input into the CNN model, and the output of the CNN model is input to fully-connected layers whose output is scoring likelihood between 1 and 0. In our design, we prepare 11 CNNs to process 11 images of the trajectories of players and a ball separately. The CNN structure is composed of a 2D Convolution layer, a Max Pooling layer, and a Dropout layer to suppress overfitting. The activation function of three fully connected (Dense) layers, to which the output values of the CNN layer are input, is Relu (Rectified Linear Unit). The activation function for the output values is Softmax, which outputs a likelihood between 0 and 1. In the training phase, the input data is the trajectory data of each player and the ball in one attack scene with the supervised label of the success (1) or failure (0) of scoring. The order of the input image data is divided between the ball, the offensive

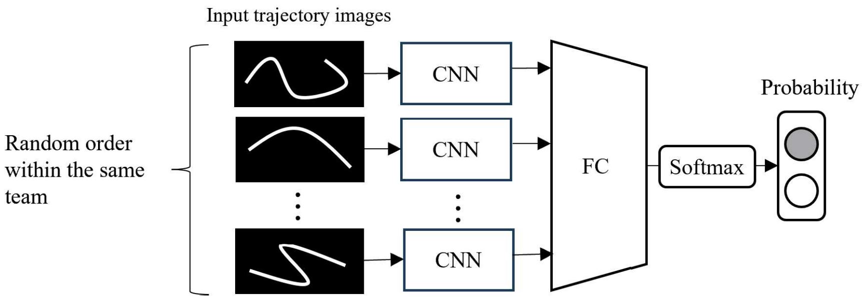


Figure 3.2: Neural networks using trajectory data of players in team sports

team, and the defensive team. We note that the order among the five players of each offensive and defensive team is random. This is because players' roles and positions dynamically change depending on situations, which is difficult to fix a specific order. If enough data is used for training, the parameters of the CNNs need to be learned for similar features within each of the attacking and defending teams. As for the adequacy of this number of data, if sufficient prediction accuracy is achieved, then we can say that they are trained independently of any particular player or order.

3.2.3 Calculation of Individual Scoring Contribution

As we described in the previous section, the scoring contribution is defined as Eq. (3.3). In this paper, only the scoring contribution is the evaluation target, so the target data is limited to scenes predicted to be successful in scoring, as shown in figure 3.3. In addition, the scoring contributions are calculated for players of the attacking team only since we focus on scoring. The scoring contribution is the amount by which the output value decreases when the input data of the target player is excluded. In other words, players whose output values decrease more significantly shall be considered to have a higher contribution to scoring, while players whose output values decrease less or remain the same shall be considered to have a lower contribution to scoring. The calculation procedure is as follows. When calculating the data for one match, the data for the attacking scenes are calculated in sequence. A single attacking scene is defined as a turnover in which the ball possession is switched from one turnover to the next, and turnover opportunities include steals, ball outs, and fouls, as well as scoring, as shown in figure 3.4. Turnover opportunities include steals, ball outs, fouls, and points. Although players frequently change during a basketball game, players are only replaced during breaks in the game, including turnovers, thus guaranteeing that no players are replaced during a single offensive scene. For each attacking scene, the output values are first calculated for all players and ball trajectory images. Then, we calculate the scoring contributions of each player in the offensive team. For the analysis of one match, we aggregate the scoring contributions of player i in all attacking scenes by i's team. We can also analyze each player's scoring contributions over a season by aggregating and averaging the player's scoring contributions of all games in the season.

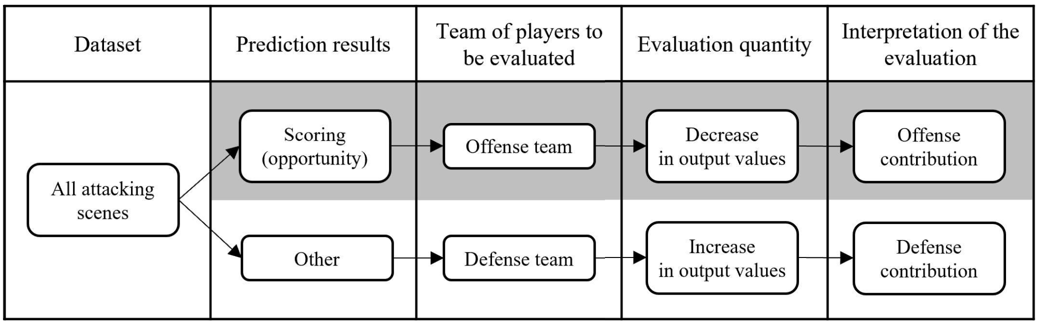


Figure 3.3: Estimation of players' contribution to a team performance using our proposed method. (In this paper, offense contribution in the upper part of the figure is the subject of the evaluation.)

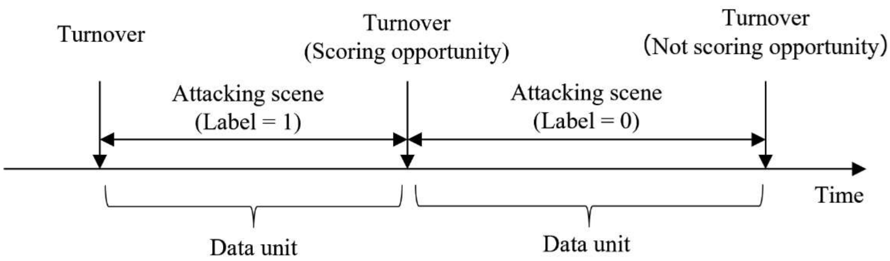


Figure 3.4: Data unit and labeling of attacking scenes

3.2.4 Visualization Method for Plays that Contributed to Scoring with CAM Applied

We developed a method that applies Score-CAM, a method that evaluates the parts that the model focuses on by using the change in output values when the image data input to the CNN is partially masked as an evaluation of the contribution of each player's specific play in the proposed method. Various CAMs have been proposed as methods for obtaining the basis for decision-making in deep learning models that take images as input. Early CAMs were limited to specific network structures, and it was pointed out that the results were unstable due to the use of back-propagation of gradients. Methods that use perturbation, such as Score-CAM, are versatile and have few restrictions on network structure, etc., and because they do not use gradients and only use forward propagation, and so it are reported to be highly stable and accurate. As shown in the conceptual diagram in figure 3.5, it consists of two steps. In Phase 1, the activation maps of the convolutional layers are obtained when the input image is input to a deep learning model that includes a CNN. When obtaining the maps, all the activation maps are obtained from the convolutional layer closest to the output. The multiple activation maps obtained are upsampled to return them to the original image size, and these are used as masks respectively. The reason for upsampling is that the size of the activation map is reduced by

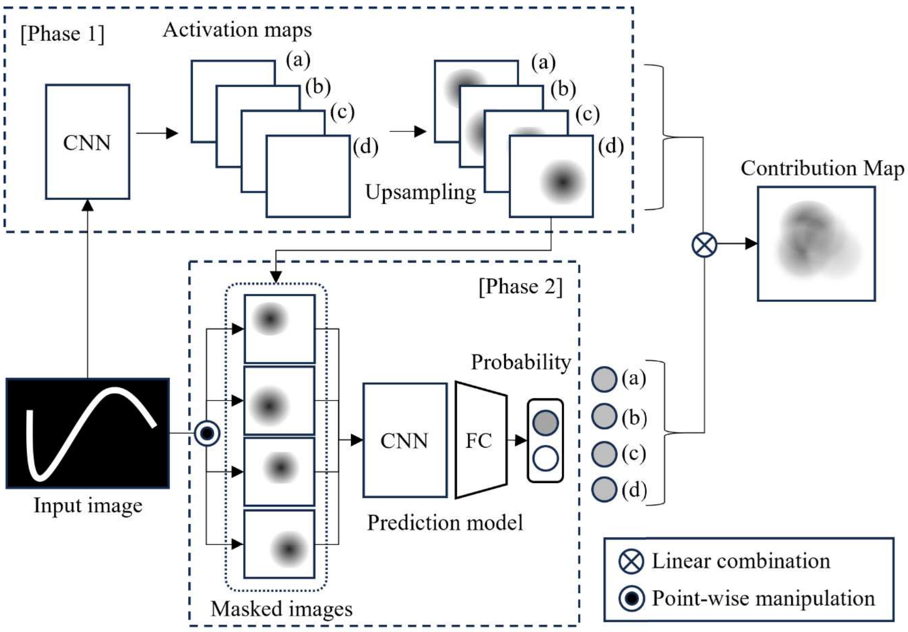


Figure 3.5: Evaluation method of contribution part of a trajectory

convolution. In Phase 2, the masks are multiplied with the input image to generate the masked input images for each activation map. The output values when the masked input images are input, that is, the probability values of the classification, are subtracted from the probability values of the original input images. The difference in probability values for each masked input image is added to the map multiplied by each mask, and the sum of all the maps is used as a map that shows the contribution to the classification result. By applying this method to a model for discriminating between playing scenes using image data of the players' movement trajectories as input, it is possible to visualize, in a way similar to a heat map, which parts of the movement trajectory the model is focusing on when decision-making. In other words, based on the information in the heat map, it is possible to identify which positions on the field contributed to scoring opportunities, etc.

3.3 Dataset

3.3.1 Tracking Data

Basketball Dataset

We use the official tracking data of the NBA, the USA's professional basketball league, from the 2015-2016 season, which has been used in various studies on basketball [13, 60]. The data is recorded by the SportsVU [61] camera system, which extracts the trajectories of players and a ball by image recognition from the videos captured by multiple cameras. The data sampling rate is 30 Hz, consisting of 2D positions of 10 players and a ball covering an area of 94 feet * 50 feet of the basketball court for 663 regular season games, excluding playoffs and other events. Each player's and ball data is distinguished by a unique ID and can be linked to event data such as shots and fouls. The event data is recorded with the tracking data and includes events such as fouls, ball outs, and steals that can be used to determine scores and turnovers with timestamps. In this paper, only attack scenes longer than 1 second were used and extracted, resulting in 102,452 attack scenes, 50,915 of which were successfully scored, as shown in table 3.1. In addition, 80% of these were used as training data and 20% as test data.

Table 3.1: Number of data in the basketball dataset

	Label		
	Scoring (1)	Other (0)	Total
Training	41,275	40,686	81,961
Test	10,262	10,229	20,491
Total	51,537	50,915	102,452

Soccer Dataset

We used tracking data from the J1 League to create data for learning and evaluation. In the J-League, Japan's men's professional soccer league (J1 is the first division of the J-League), all official matches are currently recorded by a tracking system based on video analysis using multiple cameras operated by Data Stadium Inc.. The recorded data consists of positional coordinate data of players and balls and event data. The position coordinates are time-series data recorded at a sampling rate of 10 Hz, and the number of players participating in each sample and their position coordinates are recorded. However, only the position of the ball differs from the actual trajectory, and the trajectory data is a line connecting the points where players touched the ball. The event data records the time the event occurred and the number of players who participated. The attack scene is delimited by the timing of the turnover extracted from the event data. In addition, because scoring is rare in soccer, unlike basketball, the scoring opportunity to be predicted is the entry to the 30-meter line. The 30-meter line is a line 30 meters away from the goal line, which divides the entire field into almost three sections, and the entry to the 30-meter line in the opponent's territory An opponent's entry to the 30-meter line is considered an opportunity to score a goal. When separating attacking scenes, attacking scenes involving entry into the 30-meter line were separated by one play prior to the 30-meter entry so that the entry into the 30-meter line could not be easily identified from the position of the ball or players. The study used tracking data from 50 games in the 2022 season, which were deemed suitable for the study because of the balance of teams and time periods, the relatively high number of 30-meter line entries, and the absence of foul ejections. We defined a valid play as a scene of 100 frames (approximately 10 seconds) or longer, resulting in 4325 attacking scenes, of which 1874 were 30-meter line entries, as

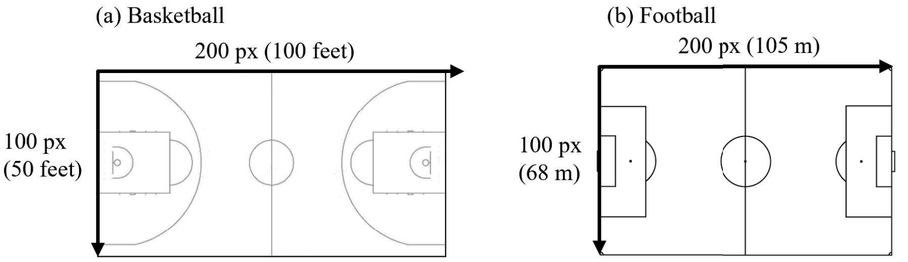


Figure 3.6: Correspondence of trajectory images to basketball and soccer courts

shown in table 3.2. As in basketball, 80% of these were used as training data and 20% as test data.

Table 3.2: Number of data in the soccer dataset

	Label		
	Scoring (1)	Other (0)	Total
Training	1,501	1,959	3,460
Test	373	492	865
Total	1,875	2,451	4,325

3.3.2 Generating Trajectory Images with Labels

The size of the input trajectory image was set to 100*200 [px*px], and the position coordinates of the tracking data acquired for the basketball court and soccer court were converted to match the input image size, as shown in figure 3.6. An example of the drawn trajectory image is shown in figure 3.7. The image is a 256-level grayscale with a black background. In other words, without the trajectory, the input would be black. The trajectory in the image is represented by dots of radius 2 [px] plotted at the positions of the players and the ball in each frame. Therefore, even if, for some reason, the recorded positions are far apart between frames, no interpolation is performed by connecting them with line segments. In addition, in order to take into account time series information and to emphasize the latest data, the brightness of the dots is represented by a gradation with 0.5 for the oldest data and 1.0 for the newest data. In basketball and soccer, both teams attack in different directions. Since the direction changes with each quarter or half of the game, the image of the attacking scene was rotated so that the direction of the attack is the same in all scenes, regardless of the team or quarter.

Figure 3.7: Example of trajectory image

3.4 Evaluation

3.4.1 Training Predictive Models

Basketball Dataset

Figure 3.8 shows our design of the team performance predictor, which predicts success or failure in scoring for the input of a 2D image representing the trajectories of all players and a ball. The CNN structure is composed of a 2D Convolution layer, a Max Pooling layer, and a Dropout layer to suppress overfitting. The activation function of three fully connected (Dense) layers, to which the output values of the CNN layer are input, is Relu (Rectified Linear Unit). The activation function for the output values is Softmax, which outputs a likelihood between 0 and 1. In the training phase, the input data is the trajectory data of each player and the ball in one attack scene with the supervised label of the success (1) or failure (0) of scoring. The order of the input image data is divided between the ball, the offensive team, and the defensive team, and the order among the five players of each offensive and defensive team is random. Of the dataset, 20% were pre-separated for testing and 80% for training, with a validation size of 10% during training. The training batch size was optimized to be 8, with a maximum of 100 epochs, and stopped when the training loss due to error backpropagation stopped decreasing.

Soccer Dataset

We trained a deep learning model that generates a discriminative model by learning a dataset of player and ball trajectories tagged with the presence or absence of scoring opportunities (i.e., whether or not a player enters the 30-meter line). The neural network as shown in figure 3.9 follows the model that discriminates scoring opportunities in basketball and is mainly based on the number of input players and the associated parameter adjustments. The input data consists of 22 players and ball trajectory image data. Since the role of the goalkeeper (GK) is clearly different from that of other field players (FPs) in soccer, the input position of the goalkeeper was fixed within a team, and the input positions of the other players were randomized. In addition, one additional fully connected layer was added

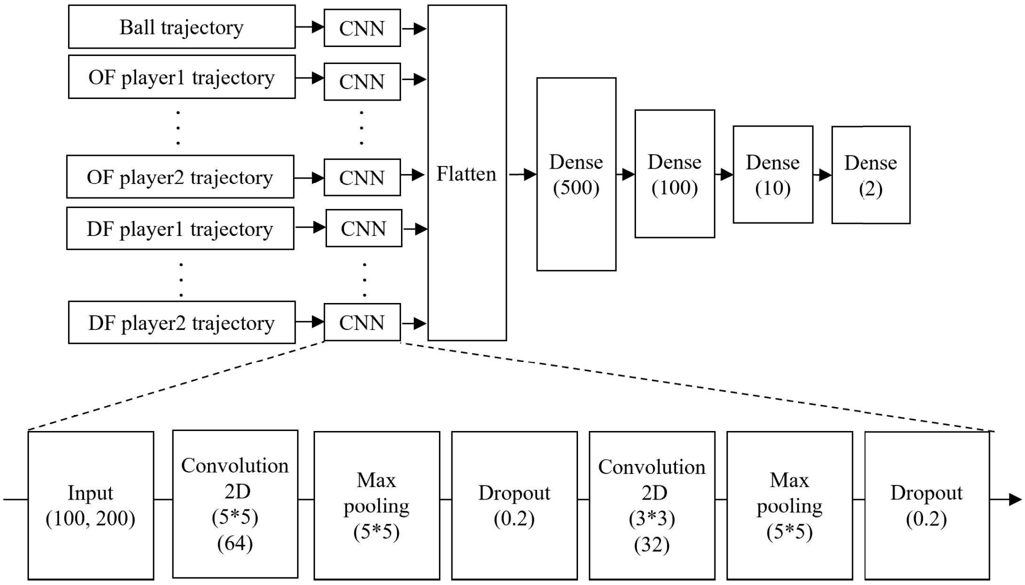


Figure 3.8: Neural network structure for predicting scoring opportunities in basketball

as the amount of input data increased, and dropout was increased to suppress over-learning. During training, the label with input data was set to 1/0, which indicates the presence or absence of the tagged 30-meter entry, and the model was trained as a binary classification model. Of the input data, 20% were pre-separated for testing and 80% for training, with a validation size of 10% during training. The training batch size was optimized to be 8, with a maximum of 100 epochs, and stopped when the training loss due to error backpropagation stopped decreasing.

3.4.2 Accuracy of Scoring Opportunity Prediction

Basketball Dataset

The results of the evaluation of the prediction accuracy of the deep learning model for successful scoring showed that the prediction accuracy was high, with an AUC of 0.927 and both macro precision and macro recall of 0.847 using all training data, as shown, as shown in figure 3.10. The Area Under the Curve (AUC) is a widely used metric in machine learning and information science, particularly for evaluating the performance of binary classification models. It represents the degree or measure of separability achieved by the model, quantifying how well the model can distinguish between two classes. The AUC specifically refers to the area under the Receiver Operating Characteristic (ROC) curve. The ROC curve is a graphical representation that illustrates the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) across different threshold settings. By calculating the area under this curve, the AUC provides a single scalar value that summarizes the model's performance. AUC values range from 0 to 1. An AUC of 0.5 indicates no discriminative

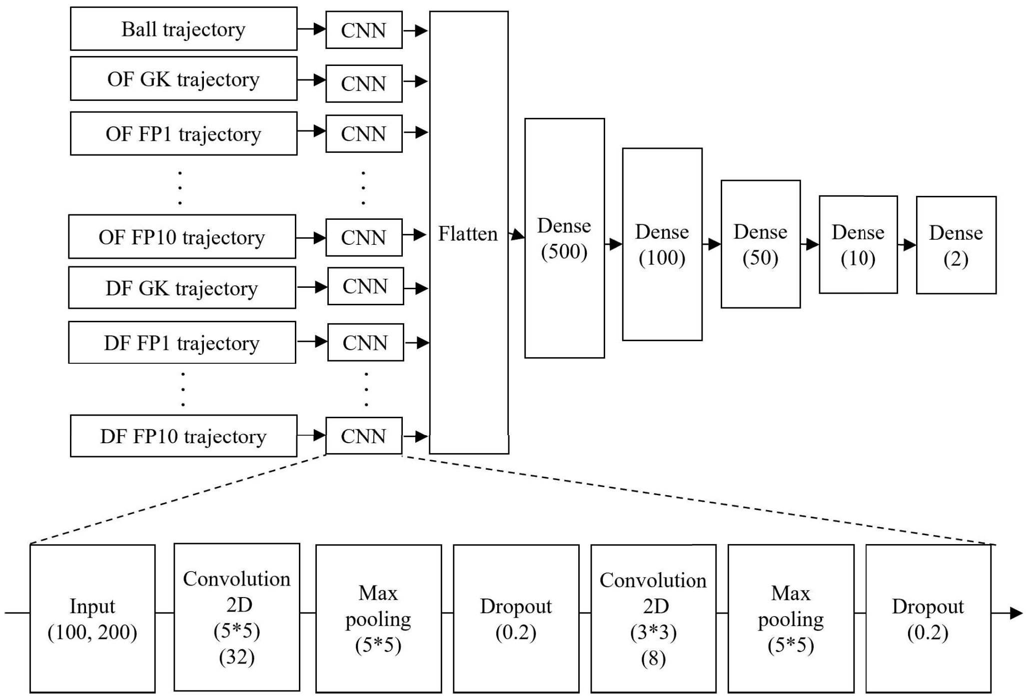


Figure 3.9: Neural network structure for predicting scoring opportunities in soccer

power, equivalent to random guessing. An AUC closer to 1.0 signifies excellent model performance, with the model effectively distinguishing between positive and negative classes. An AUC below 0.5 suggests a model performing worse than random, which might indicate an issue with the model or data. The advantage of using AUC is its ability to provide a threshold-independent performance measure, making it a robust metric for comparing models, especially in imbalanced datasets. It helps researchers and practitioners understand the overall effectiveness of a model in identifying true positives while minimizing false positives.The accuracy was saturated when the training data was above 40,000, indicating that the model was sufficiently trained by deep learning. In the label-by-label evaluation of scoring success or failure, as shown in table 3.3, under the condition that the number of test data for each label was the same, precision and recall were 0.867 and 0.820 for label 0 and 0.827 and 0.873 for label 1, respectively. This showed that the differences between labels were small, although the precision of label 0 tended to be slightly higher and the recall of label one slightly higher. We can say that we have obtained a performance prediction model that performs well enough on the basketball data set.

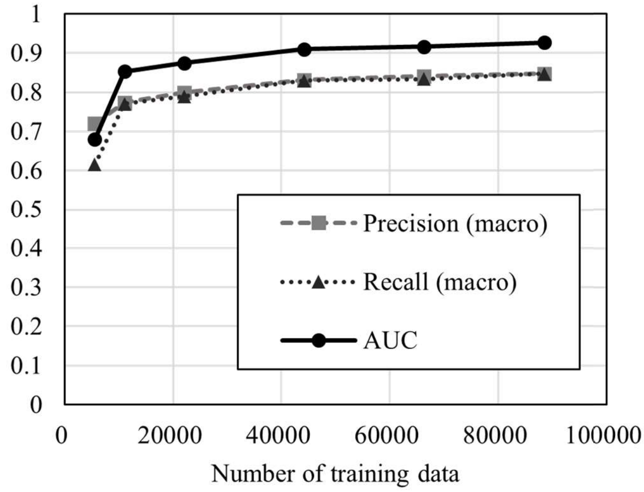


Figure 3.10: Relationship between the number of training data and prediction accuracy

Table 3.3: Predictive performance of the basketball dataset

Label	Precision	Recall	AUC
1	0.827	0.873	0.927
0	0.867	0.820	

Soccer Dataset

The prediction model was trained using training data from the dataset of movement trajectories and labels for success or failure of the 30-m line entry shown in table 3.2, and its performance was checked using test data. The accuracy of each label was calculated and summarized in table 3.4. As shown in the table, the AUC is 0.872 and exceeded 0.8, which is lower than the accuracy of the prediction model for basketball scoring opportunities but still high by general standards. Improving the prediction accuracy requires increasing the number of data, which is less than in the aforementioned basketball. Furthermore, looking at the prediction accuracy for each label, the recall for label 1, indicating 30m line entry, was 0.693, which is lower than 0.825 for label 0 and is not balanced against the close values of 0.776 and 0.762 for precision, respectively. This could be attributed to the different number of training data for each label, as shown in table 3.2. To remedy this, we also performed training with the number of labels aligned with label one by undersampling the data for label 0, but no significant improvement was observed. It is possible that the difficulty and properties of the prediction may differ from those of basketball and should be verified in future evaluations with more data.

Table 3.4: Predictive performance of the soccer dataset

Label	Precision	Recall	AUC
1	0.776	0.693	0.872
0	0.762	0.825	

3.4.3 Validity of player's contribution

Basketball Dataset

To evaluate each player's scoring contribution, we obtained performance data from the official NBA website for comparison. Typical individual player performance includes values that evaluate offense, such as the number of points and assists, and defense, such as the number of blocks and steals. These evaluation values are calculated for each game and throughout the season. Efficiency (EFF) [62] is commonly used as the overall metric, considering offensive and defensive values and negative values such as missed shots, etc. EFF is calculated as:

$$\begin{aligned} EFF = &(Points + Rebounds + Assists + Steals + Blocks) \\ &- \{(FieldGoalsAtt. - FieldGoalsMade) \\ &+ (FreeThrowsAtt. - FreeThrowsMade) + Turnovers)\} \end{aligned} \tag{3.4}$$

Rebound is the number of times the ball is grabbed or played toward a friendly player when a shot is missed. Field Goals Att. is the number of shots attempted during the game, and Field Goals Made is the number of successes. Free Throws Att. is the number of scoring opportunities obtained by fouls and violations by the opposing team, and Free Throws Made is the number of successful attempts. This turnover is the number of times the ball was lost. In this study, EFF for one season was used as the overall player metric for the period for which the data under study was measured. The Speaman's rank correlation coefficient was used to evaluate the correlation between the score contribution and Efficiency for each player. In evaluating the correlation, a correlation coefficient of 0.3 or higher was considered to be correlated using standard criteria, and the significance level was set at 0.05. The average of the contribution to points per game for the season and the average of the Efficiency per game for the season were used for the comparison. As a result of limiting the number of players who played, 143 players were included in the study. We also evaluated and compared the correlation coefficients with the number of Points, Assists, and Rebounds per game, representing scores related to an offense other than Efficiency. The results for the subject season were obtained from the NBA's official website. In evaluating players' scoring contribution, the rank correlation between scoring contribution and Efficiency was calculated using the scoring contribution ordered by Efficiency ranking shown in figure 3.11. The correlation coefficient with Efficiency was 0.373. The p-value was less than 0.001, below the statistical significance level, indicating a significant correlation. The figure also shows the scoring contribution sorted by the ranks of other representative offensive indicators, i.e., Points, Assists, and Rebounds. The correlation coefficients of 0.168 and 0.127 with the rankings of points and assists were below the threshold for correlation, and the p-values were above the significance level. For rebounds, the coefficient was 0.272, below the criterion, and the p-value was below the significance level. In other words, these single scores were not significantly correlated. Table 3.5 also summarises the rank correlation coefficients compared to non-offensive indicators, running distance, and speed. Similarly, no indicators with a correlation coefficient of 0.3 or higher were found. This means that players with high scoring contributions for the proposed method tend to be similar to players with high overall evaluation metrics such as EFF, rather than having high specific offensive metrics. In addition, the longer trajectories did not lead to higher score contributions, indicating that they were independent of the magnitude of the input values. The evaluation results indicate that in basketball, players' trajectory data contain sufficient information related to the success or failure of scoring. In previous studies, various indicators such as shooting position and angle have been used by trial and error to extract plays related to scoring [63, 64]. The group trajectory data contains information on those indicators based on prior knowledge, suggesting that the proposed method can be used for highly accurate prediction without feature extraction. On the other hand, the proposed method can not explain what kind of play contributes to scoring goals in attacking scenes, which is a future issue. In the future, it will be possible to extract the team and individual plays that contributed to scoring by analyzing patterns by collecting and clustering trajectory data that were predicted to score, and by analyzing changes in predictions by extracting or excluding only a part of the time in an attacking scene. In contrast to the conventional score, the player's contribution to scoring in the proposed method considers all movements in each attacking scene. It is a comprehensive evaluation index similar to Efficiency. In conventional scores, only those directly reflected in numerical values, such as goals and assists, are considered, and other contributions are not considered. In this point, Efficiency combines scores to provide a more comprehensive evaluation. Still, it does not reflect the quality of play that is not reflected in the score, and it is also pointed out that the stronger the team, the higher the value tends to be. In recent years, Player Efficiency Rating (PER) has been proposed as an evaluation index to improve such problems, but it is still a value calculated based on scores [65]. Our proposed method of scoring contribution is not an addition of scores, so it is not affected by the strength of the team, and a proposal can be obtained that fair evaluation is possible. In other words, it contributes to reevaluating and scouting players who have not been adequately evaluated in the past and determining rewards through fair evaluation.

Soccer Dataset

We compared the contribution of players calculated by the proposed method with other indices of players' contribution to the attack created by other experts based on their experience. However, since the contribution calculated by the proposed method is a completely new indicator that does not use

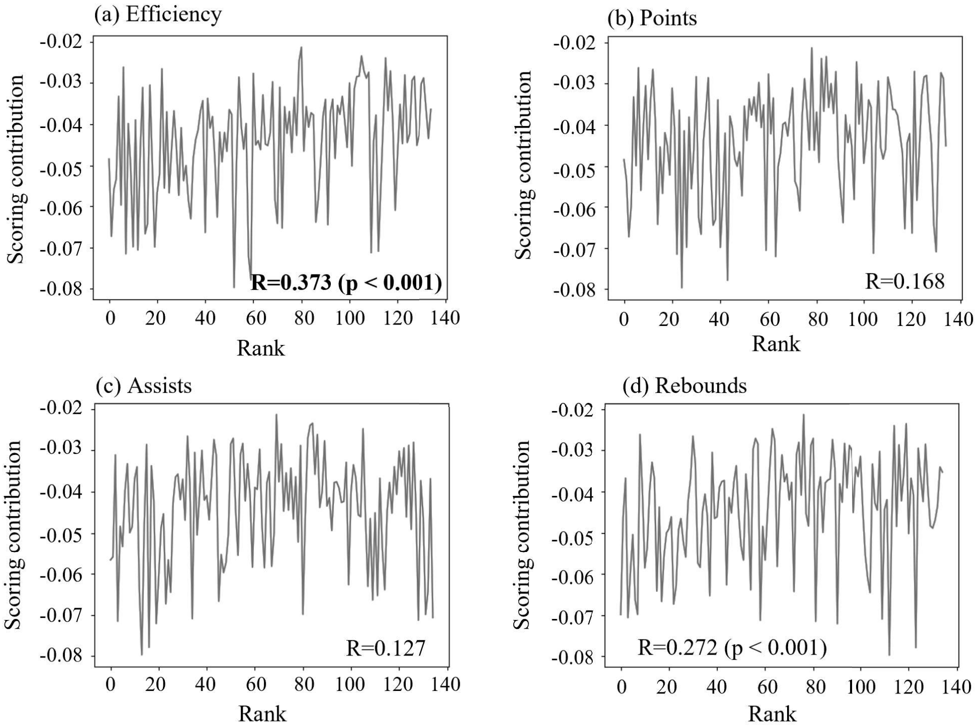


Figure 3.11: Offense contributions sorted by annual ranking of Efficiency, Points, Assists, and Rebouds averaged per game and Pearson's rank correlation coefficients

human hypotheses, it does not match existing indicators, and we expect it to provide new value that does not rely on human experience. As a comparison, we used CBP (Chance Building Point) as a contribution to scoring opportunities, referring to Football LAB [66], an information provision site by Data Stadium, Inc. that provides original analysis results, etc. CBP is a measure of the "play items (passes, passes, etc.)" that occur on a soccer court and is used as a measure of the "score" of a player's game. CBP defines the plays that occur on a soccer court in terms of "play items (passing, dribbling, crossing, etc.)" and "areas" and scores them using a calculation formula based on two perspectives: how much the play leads to a shot, and how difficult the play is. Since the reachability and difficulty of a shot differ from area to area, the score is calculated based on how much each player's play contributed to the shooting chance, taking into account each of these factors. At the same time, we also compared the proposed player contributions to basic offensive indicators such as shots and passes per minute played. Considering the small number of 50 soccer games used in this study and the biased nature of the teams' results, we compared the contribution of each player in attacking scenes within a team and compared it with the existing indices by rank correlation coefficients. The target players were those who had played at least one game. Using the proposed method, we calculated the rank

Table 3.5: Rank correlation coefficients between offense contribution and existing indicators ($* : p - value < 0.001$)

Target Indicators	Correlation coefficient
Efficiency	0.373*
Points	0.168
Assists	0.127
Rebound	0.272*
Steals	0.044
Blocks	0.297*
Turnover	0.172
Distance (offense)	0.027
Speed (offense)	0.145

correlation coefficients between offensive team players' contributions and Pearson's rank correlation coefficients of the existing indices, calculated using the soccer dataset as input. Figure 3.12 shows the rank correlation coefficients with CBP. We were able to confirm a correlation of 0.3 or higher for 7 of the total 18 teams. Therefore, although the contribution index was calculated using data from a short period of time, it shows a similar trend to the annual contribution index to scoring opportunities analyzed by experts. Next, figure 3.13 shows the correlation coefficients with the annual frequency of shots, dribbles, passes, and pass receptions. In particular, about ten teams show a correlation of 0.3 or higher for shots, dribbles, and pass receptions. From these correlations, it can be said that the contribution of the proposed method represents the characteristics of offensive play of players, although it was calculated from short-term data. In addition, since it is calculated from movement trajectories, it may reflect the characteristics of receiving passes more strongly than the characteristics of making passes. Thus, it is different from existing performance indicators, although similar in some aspects, and can be said to indicate performance that can be evaluated solely from movement trajectories. However, since the soccer tracking data that could be used in this study was limited to 50 games, and the correlations were not all statistically significant, continued evaluation with additional data is needed for more accurate validation. Finally, we checked the map that visualizes the contribution of each player's play, and an example is shown in figure 3.14. The figure shows a map that combines a heat map of overall contribution with a trajectory image. We were able to confirm that it is possible to output a heat map where areas of high contribution are red and areas of low contribution are blue. In the figure, we can see areas of high contribution within the trajectory, and we can also see areas that are red even in cases where there is no trajectory, which is rare. In this way, we were able to confirm that the fact that there is no trajectory in certain areas also contributes to the identification.

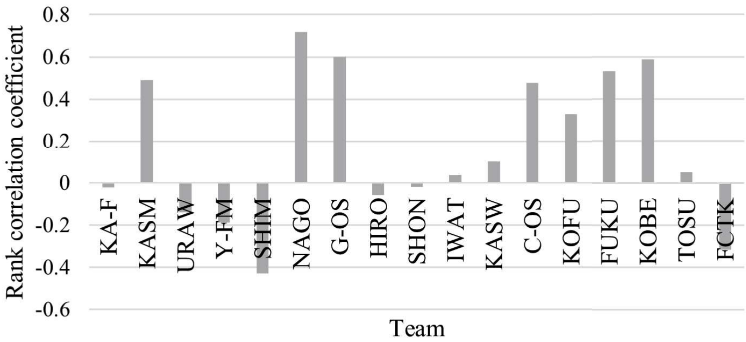


Figure 3.12: Rank correlation coefficients between offensive contribution and CBP for the J1 soccer teams

3.5 Discussion

3.5.1 Application

The results of this research can be applied to the real-time prediction of scores and score-ness during basketball and soccer games, evaluation of player contributions without using scores, and training and team tactics. Recently, various sensor data and tracking data have begun to be used for training [67]. In training applications, the learned model can be used to evaluate movements during practice and provide feedback to increase the predictive value of scoring and contribution. As a specific feedback method, it is being considered to provide coaches and players with the reasons for high-scoring performance evaluations based on trajectory data, as well as video footage, and to express situations in a more intuitive way for players using VR and AR [68]. Through this feedback, it is possible to identify patterns of movement that contribute to scoring that have not been noticed by players or coaches before. For use in team tactics, the learned model can be used to search for combinations of players and formations that will increase the predictive value of scoring the most points. Furthermore, quantitative evaluation of the proposed method can also be used in player recruitment to find players who are in good shape or have developed in a way that the team's coaches missed. A new application is a possibility of automating the labeling of data for different purposes of machine learning. Until now, finding good moves that lead to scoring opportunities, excluding actual scoring scenes, required much effort by a person with the expertise to visually check and label videos and other data. By further developing the proposed method, the labeling of scenes can be automated, and the generation of machine-learning models can be made much more efficient. As a developmental use of the highly accurate score prediction model, it is also expected to be used for Reinforcement Learning and Generative Adversarial Networks [69] that learn and automatically generate new ways of player movement that have never been seen before. The proposed method is also expected to be applied to sports other than basketball. In team sports other

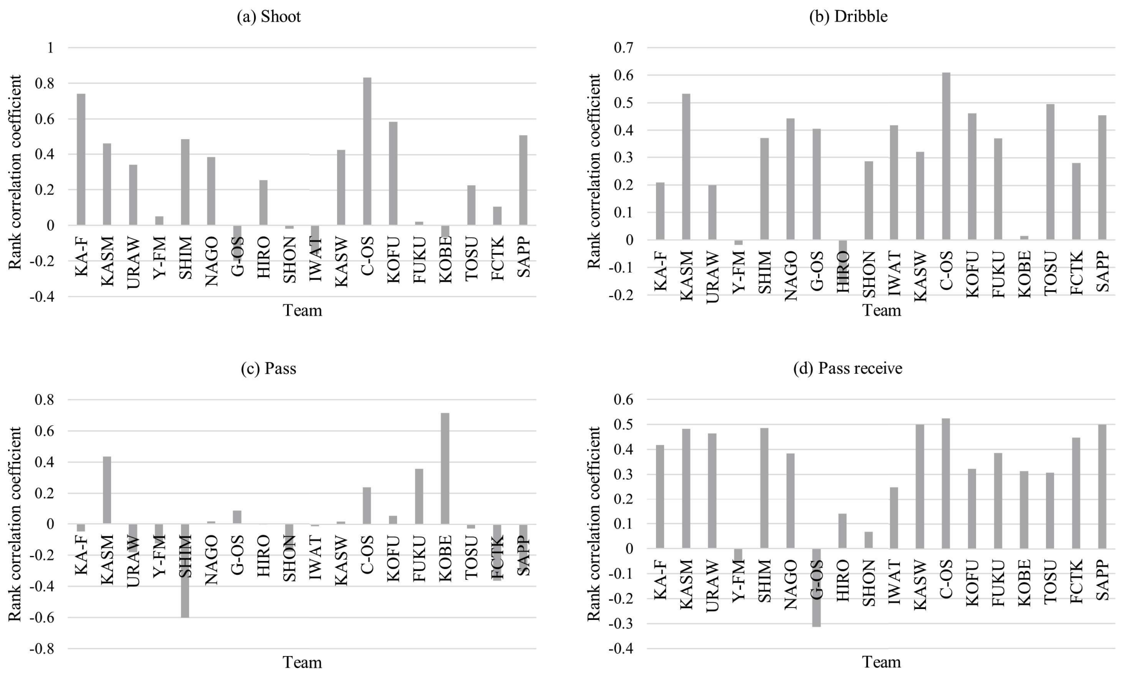


Figure 3.13: Rank correlation coefficients between offensive contribution and existing annual play frequencies for the J1 soccer teams (the team names are abbreviated).

than basketball and soccer, tracking data using video and sensors has become common in popular professional sports such as soccer, and tracking data has been accumulated [7]. In addition, some professional soccer TV broadcasts have already introduced AI-based prediction [70]. The proposed method can also learn attacking scenes in other sports using existing data. On the other hand, while basketball did not need to be considered because of the large number of scoring opportunities, other sports with fewer scoring opportunities, such as soccer, need to consider optimal teacher data, such as using labels for states that are close to scoring other than scoring.

3.5.2 Limitation

Since the goal's location on the basketball court is fixed, there is a strong association between scoring and the ball's trajectory. Therefore, while the trajectory to be input is affected by the interaction between players during the process of reaching the goal, based on the focus of the prediction model for the input image using CAM, it is thought that the positioning and movement near the goal strongly contribute to the prediction of points and contribution to points. Based on this, in order to evaluate the quality of the movement of individual players and groups of players and the process of reaching the goal in more detail, it is necessary to divide the time series and analyze the trajectory near the shot and other trajectories separately. In this respect, we will consider in detail in Chapter 5 the application of

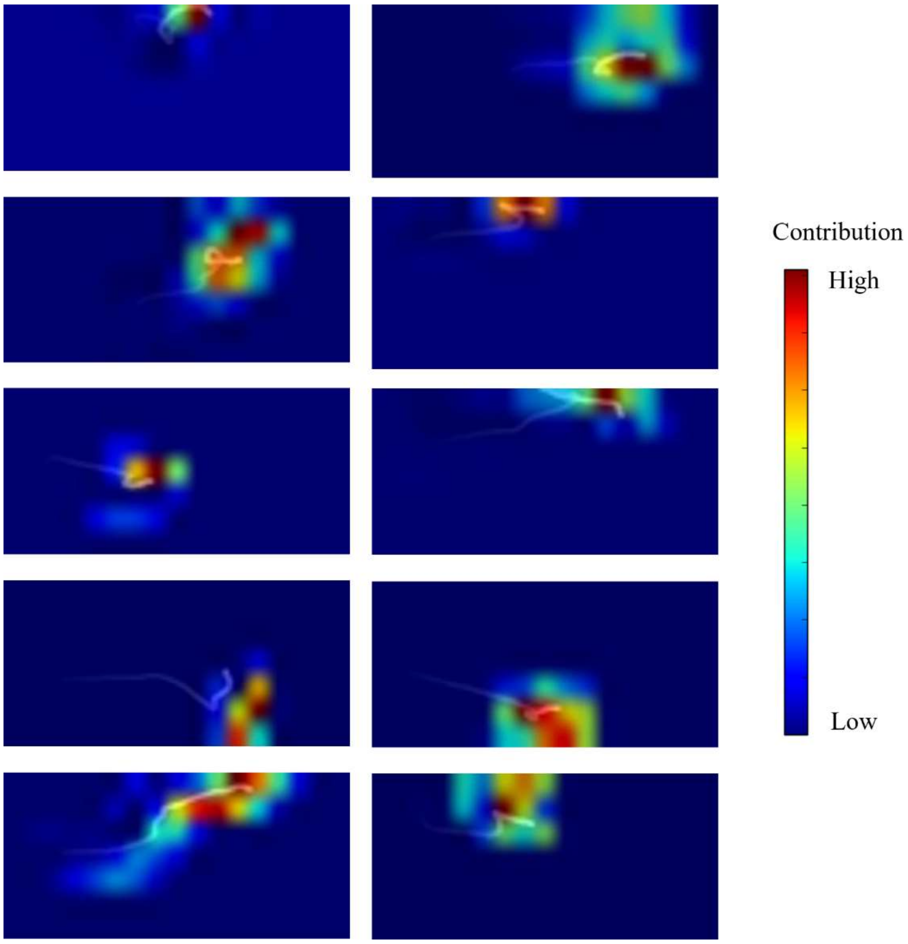


Figure 3.14: Contribution maps of the offense field players when entering the 30-meter line

deep learning, such as RNNs that learn time series data, and neural network structures that take into account the interactions between players at each point in time. In addition, the proposed method only considers the contribution of a single player to the score and does not consider combinations of multiple players. In reality, combinations of two or three players also contribute to the score, so it is necessary to evaluate the decrease in the prediction value when multiple players and single inputs are excluded in the future. In addition, the proposed method's scoring contribution considers only one person's contribution and does not consider combinations of multiple players. In reality, the combination plays of two or three players also contribute to scoring goals, so it is necessary to evaluate in the future the decrease in predictive value when excluding the input of multiple players and one person. As a method for extracting good sporting performances, it is debatable whether all machine learning answers, such as deep learning, are correct. A limitation of this paper is that only a short period of data was trained.

Since only data from 50 games in one season for soccer and one season for basketball were trained, only the trends of scenes scored in that season are reflected. Since sports tactics are ever-evolving, new data and data from different leagues must be constantly added to the learning process to make it practical. Even for high-level players, it is unknown whether their learning data will be effective for game tactics and player development at other levels and generations. In the future, it will also be necessary to change the quality of the learning data depending on the application. There is also a need to change the quality of the training data depending on the application in the future. Furthermore, machine learning is sensitive to statistically dominant data and may not be suitable for discovering 'exceptionally good plays.' Machine learning models are based on past dominant good results, making it difficult to predict or extract new results beyond the past. In other words, the proposed method does not reflect creativity, which is an important factor in sports. At present, it is considered that the proposed method can lead to the evaluation and extraction of probabilistically good plays from the perspective of the period for which the learning data was obtained.

3.5.3 Future Work

There is a need for a more developed understanding of the factors that contribute to performance, such as scoring in team sports, than the individual contribution proposed in this study. Collaboration by specific pairs or combinations of several people may have a greater impact on performance than individuals, and research that extends the assessment of individual contributions is possible. The impact of specific plays, rather than on a person-by-person basis, is also important for understanding behavior. In addition to individual explainability, the ability to explain the time and place of the contribution to performance and the coordination with other players, as in the case of image recognition XAI, can contribute more to the development of sports training and strategy. Improved explainability can also contribute to a better understanding of machine learning models and the development of new technologies. For this reason, future research is being considered to evaluate the positional relationships of trajectories and the causality of movements between balls and players, which could deepen our understanding of group behavior. An approach that predicts results from earlier stage data is also needed for developmental applications. The proposed method more strongly reflects movements in the time period close to scoring. If the probability of scoring could be predicted from a time as far in advance of scoring as possible, it would be possible to construct a more varied strategy for the process leading up to scoring, thus contributing to the development of collective sports.

3.6 Conclusion

We proposed a deep learning network structure in which multiple trajectory images are input to different CNNs in order to predict scores from trajectory data of players and balls in team sports and to estimate the contribution of players to scores. We evaluated our system on real professional basketball and soccer data. Accuracy evaluation results showed high prediction accuracy with AUC of 0.927 and 0.872, respectively. In addition, we proposed a method that uses the decrease in predicted scores when the input data of the target player is excluded as the player's scoring contribution. The calculated scoring contribution was significantly correlated with the annual Efficiency, which is used as an existing overall indicator in basketball, and was not correlated with typical offensive scores such as Points, Assists, and Rebounds. We conducted a similar evaluation in soccer and found correlations within multiple teams in CBP, an existing overall offensive measure, and in the number of shots, dribbles, and pass receptions. The evaluation results indicate that the proposed player contribution to scoring opportunities can be a new evaluation indicator that reflects the offensive characteristics of players as manifested in their movement trajectories. It is expected to be applied to team sports such as basketball and soccer games and training.

Chapter 4

Evaluation of Cooperation Graphs between Players in Team Sports using Transfer Entropy

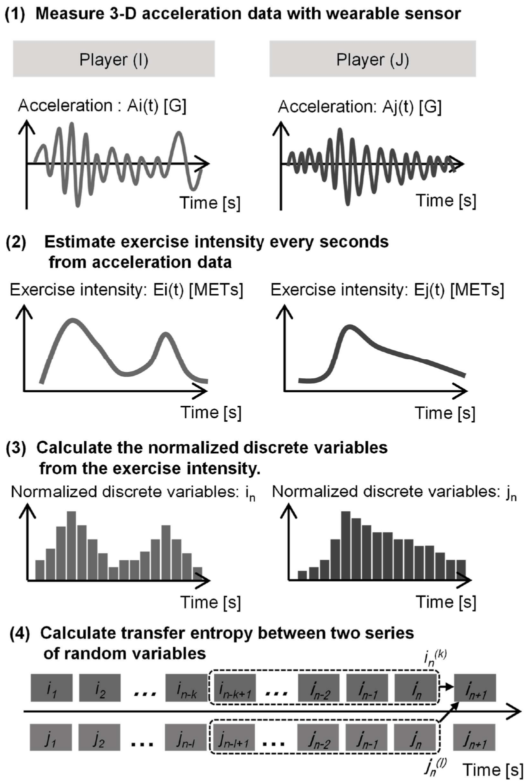
4.1 Introduction

With the strengthening of training for Olympic athletes in each country and the expansion of investment in professional sports, analysis using IT in the sports field is attracting attention. In particular, in group sports such as soccer, where professional sports are well developed, wearable devices equipped with GPS and acceleration sensors are becoming widespread, and various sensor data are being used to manage the training intensity and condition of athletes [71, 72]. In recent years, with the development of video analysis technology, technology to recognize athletes and their movements from high-definition video is being researched and developed, and in the future, non-contact sensing that places less of a burden on athletes is expected. At present, since the identification of individuals is still done using their uniform numbers and manual correction is also necessary, this technology is mainly used for professional sports matches held in stadiums where overhead video can be obtained, and wearable devices are widely used for measuring daily training and practice matches. In order to improve major team sports such as soccer, it is important to evaluate performance based on quantitative data [73, 74]. Team sports performance can be divided into individual performance and organizational performance, and to evaluate individual performance, methods are used to evaluate physical performance such as distance moved, speed of movement, and exercise load using data from wearable devices and video analysis, and this is also used to prevent injuries [75, 76]. On the other hand, with regard to the organizational performance is also important in team sports, research has been carried out on analytical indicators for analyzing the coordination between players involved with the ball and their tactics with the opposing team, in order to evaluate whether the team's tactics are functioning as expected, and mathematical approaches to the distances and positional relationships between players and the goal have been presented [77, 78]. The coordination and tactics between players are based on communication with players on the same team based on individual visual information and on predictions of the next play by the opposing team and can be seen as a form of latent information transmission through the senses of the players. This kind of latent information transfer is thought to be taking place constantly throughout the course of a sporting match, and in ball games such as soccer, it is necessary to pay attention not only to scenes involving the ball but also to scenes not involving the ball. Many coaching theories advocate the importance of movement and coaching methods in scenes where the ball is not in play [79, 80, 81]. However, previous research has not analyzed the consistent communication between players, including scenes where the ball is not concerned. In addition, many analyses have been carried out visually using video, and it has been difficult to confirm the effects of training in real-time using the results of the analysis in the field of coaching or to grasp changes such as the improvement of players after the fact because it is not possible to confirm the movements of all players being coached by the sports coach in detail or to make quantitative comparisons. In this study, we propose an analysis method using transfer entropy with the aim of quantifying the systematic performance of players, focusing on the potential information transfer between players. Transfer entropy is an indicator that expresses the causal relationship between time series data based on information theory, and has been studied in the context of neural circuits [82], human multisensory communication [83], the mutual relationship between SNS and product purchasing [84], joint attention [85], and investment indicators [86]. In addition, since it can be applied to independent data measured simultaneously, it can be implemented in the data analysis of wearable devices, which are becoming increasingly popular. We examined an analysis method using acceleration data measured by a wearable device that can be used easily in daily training and evaluated its validity and usefulness by collecting and analyzing data on players in a soccer match, which is a team sport.

4.2 Method

4.2.1 Estimating Linkage between Players Using Transfer Entropy

In order to quantify the systematic performance of players, we examined an analysis method for potential information transfer between two people using transfer entropy. In brain measurement, which is also in the field of human measurement, moving entropy is also used as a method for estimating neural transmission from time series signals that show the activity of each neuron. Typical methods for examining the relationship between two signals include linear analysis methods such as cross-correlation and coherence functions, and nonlinear analysis methods such as mutual information and relative entropy. These are suitable for evaluating the strength of the relationship between two signals, but because they are symmetrical in shape for the two signals, they are not suitable for analyzing the direction of the flow of information, such as causality. In contrast, the application of transfer entropy has been promoted in recent years as a method suitable for estimating causal relationships, and it has



$$T(J \to I) = \sum P\left(i_{n+1}, i_n^{(k)}, j_n^{(l)}\right) \log \frac{P\left(i_{n+1} \middle| i_n^{(k)}, j_n^{(l)}\right)}{P\left(i_{n+1} \middle| i_n^{(k)}\right)}$$

$$i_n^{(k)} = (i_n, i_{n-1}, \cdots, i_{n-k+1})$$

$$j_n^{(l)} = (j_n, j_{n-1}, \cdots, j_{n-l+1})$$

Figure 4.1: Process to calculate transfer entropy between two players

been found that transfer entropy can also be applied to signals that undergo non-linear fluctuations. In this study, we propose a method for visualizing the structure of latent information transmission in team sports by calculating transfer entropy from acceleration signals of two players playing in the same field at the same time. The flow of the proposed method is shown in figure 4.1.

In the process (1), the movements of each individual are measured using an accelerometer with time-synchronised data. In the second step, the amplitude of the acceleration data is summed over a one-second time window to produce a measure of the intensity of the player's movement. The one-second time window is sufficiently fine-grained to indicate the transition between plays in field sports such as soccer, and the data can be compressed to a practical size suitable for communication and data storage. In the process described in (3), the time series data for each individual's variable (exercise intensity) is converted into a normalized discrete probability variable by converting each individual's exercise intensity into a discrete value that becomes a histogram with k divisions. The number of divisions k for the histogram was determined based on Sturges' formula. As a result, the differences in individual exercise intensity can be absorbed and expressed as variables with relative widths. In (4), the amount of information transfer between the variables of the two athletes is calculated from the transfer entropy. Here, if the elements of the probability variables I and J at time step n are i_n and j_n, the transfer entropy $T(J \to I)$, which shows the effect of J on I, is calculated using the formula in the figure.

In this study, we set $k = l = 1$ to evaluate using the simplest model that only considers the influence of the previous movement of the players. In this case, $P(i(n+1) \mid i(n), j(n))$ represents the joint probability of $i(n+1)$, $i(n)$ and $j(n)$, and $P(i(n+1) \mid i(n))$ represents the conditional probability of $i(n+1)$ when I is $i(n)$. When there are two time-series data, I and J, the degree of uncertainty that changes relatively when the past series of J is added can be measured compared to the uncertainty of predicting the next state of I from the past series of I. The transfer entropy takes a value between 0 and 1, and the larger the value, the stronger the causal relationship between I and J. In this study, we use time series data for exercise intensity aggregated every second, so we calculate the influence of exercise intensity J on I one second later, which is considered to have the strongest influence in the near future.

4.2.2 Visualization Methods for Organizational Networks

Using the transfer entropy calculated above, we will visualize the network of potential information transfer in a team playing a team sport. In the visualization, we will use a directed graph network, as shown in figure 4.2, to which various graph analysis methods can be applied. Directed graph network analysis has been applied to various problems, such as improving the efficiency of content searches on the Internet and analyzing communities in social network services (SNS) [87, 88]. The direction and presence or absence of connections in the directed graph were determined based on whether the difference in bidirectional transfer entropy between two players was above a certain threshold, indicating that there was clear information transfer in one direction between the players. The threshold P was tuned so that the average degree per player was about half the number of players, from the perspective of the amount of information that is easy to recognize when visualized as a network.

Figure 4.3 shows the results of visualizing the data from the soccer measurement experiment discussed later in this study. To make it easier to understand the state of the network, the lines of the nodes representing the players for each team are solid or dotted lines, and the nodes have the players' unique shirt numbers. The positions of the players are based on a spring model. In the example displayed in Figure 4, we have visualized the two most typical of indices commonly used to measure network characteristics: degree centrality and closeness centrality. Degree centrality indicates the number of direct connections to other nodes in the network and can be judged visually by looking at

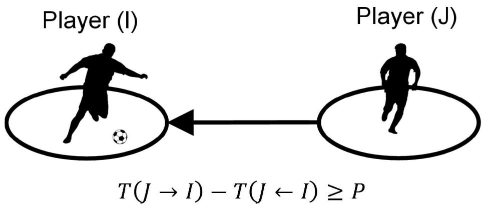


Figure 4.2: Digraph showing information transfer between two players

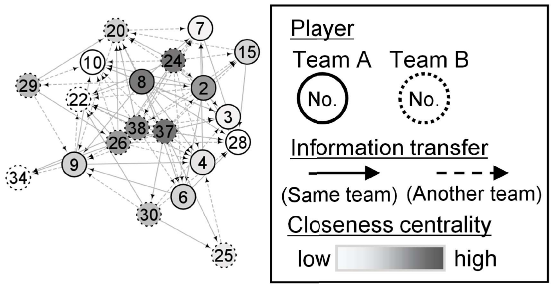


Figure 4.3: Example view of network showing information transfer

the number of arrows connected to the node (its degree), and the relative size of the degree within the whole network. For example, Team A's No. 28 and Team B's No. 22 can be seen to have a high degree of centrality, as they have many arrows connected to them, especially arrows from the opposing team. Proximity centrality is an indicator that evaluates the degree of closeness between players (vertices). In other words, players with high proximity centrality are those who are highly influential in the network as a whole, in that they are able to easily make contact with all the other players in the network and provide information quickly. Proximity centrality is expressed as the darkness of the color used to fill in the graph. For example, we can say that player number 8 from Team A and player number 37 from Team B have high proximity centrality. The calculation method is to first calculate the total of the shortest distances from one vertex to another. The shortest distance is the number of edges that must be traversed to reach a vertex from another vertex in the shortest possible time. Proximity centrality is the reciprocal of the shortest distance from one vertex to another vertex, so we calculate proximity centrality using the shortest distance calculated. Proximity centrality also differs depending on the number of vertices in the network, so we standardize it by dividing it by the value that would theoretically give the maximum proximity centrality for that group. In a group with a population of n, the maximum value of the proximity centrality is 1/(n-1), so the standardized proximity centrality is the value divided by 1/(n-1).

Figure 4.4: Image of wristband-type device

4.3 Measuring Device and Dataset

4.3.1 Measuring Movement using Wearable Devices

In order to collect data on the movements of each player, we considered using a wearable device equipped with an accelerometer. In order to collect data on the movements of each athlete, we considered using two types of wearable devices equipped with accelerometers. The first device is a wristband-type device that can easily measure the movement status of each person simply by wearing it on the arm. In addition, because it is small and lightweight, it is possible to minimize the burden and discomfort for the player wearing it. Therefore, by using an accelerometer, it is possible to measure movement data in games and during daily training, without being restricted by location. In this study, we used a wristband-type wearable device (UW-301BT manufactured by A&D Co., Ltd.) (Figure 4.4, table 4.1) after confirming that it did not interfere with the players' play in the target sport, such as soccer, and that it did not pose any safety issues. It has been reported that the acceleration data measured by the wristband-type wearable device used in this study for the hand opposite the dominant hand has a high correlation with the exercise intensity measured at the body trunk [89]. In the measurement experiment conducted in this study, each player was asked to wear a wristband-type wearable device on the wrist opposite the dominant hand, and the acceleration data of the arm was measured during the match and training. The wristband-type wearable device used in this study is equipped with Bluetooth wireless technology, but the wireless communication range is only about 10 m, so it is not possible to cover the entire sports field. Therefore, the measured acceleration data is collected via a USB connection after the match or training session has finished.

As a second device, we developed a prototype of a wearable device called the 'TS11GPS' as a unique tracking system for collecting and analysing data on the movements of athletes in real time. In developing the device, we aimed to create a small, lightweight device that would not be too much of a burden on the athletes during experiments, while also incorporating multiple types of sensors to measure movement and location information in more detail. The appearance of the TS11GPS and the

Table 4.1: Specification of wristband-type device

Feature name	Value
Size/Weigh	20 mm [W] $\times$39 mm [L] $\times$14 mm [H]/ 20 g
Measurement data (sampling rate)	Acceleration (20 hz)
Communication	Bluetooth, USB
Battery lifetime	10 days

Figure 4.5: Image of our prototype wearable device

structure of its circuit board are shown in Figure 1, and its detailed specifications are shown in table 4.2. The device is worn in a special vest with cushioning and elasticity, as shown in figure 4.5, and is fixed in place in the back pocket as indicated by the arrow, so that it does not change direction relative to the body between players. The sensors installed are a 9-axis sensor commonly used as an IMU (Inertial Measurement Unit) in human measurements, and a GPS sensor used for speed and position detection, particularly in sports tracking. The data recording frequency is the same as that of the major tracking system (Catapult Sports product) used in professional soccer, with a 9-axis sensor at 100 Hz and GPS at 10 Hz, and it is equipped with flash memory to record the measured raw data. In addition, the measured data is transmitted via 2.4 GHz band wireless, and data can be collected at 1-second intervals using the dedicated receiver and PC. The size of the device has been reduced by using low-power technology that finely controls the processor clock to keep the battery size to the same level as that of the first wristband-type device. Compared to the lightest product among the aforementioned main tracking systems, the TS11GPS weighs only 21.5 g, about half the weight. The TS11GPS makes it possible to collect detailed tracking data for research purposes without interfering with the athlete's training.

Table 4.2: Specification of wearable device

Feature name	Value
Size/Weigh	35 mm [W] ×65 mm [L] ×11 mm [H]/ 21.5 g
Measurement data	GPS (10 Hz) Accelerometer (100 Hz) Gyroscope (100 Hz) Magnetometer (100 Hz)
Communication	IEEE802.15.4, USB
Battery lifetime	3 hours

4.3.2 Collecting Data for Evaluating the basic performance of our prototype device.

In order to evaluate the basic performance of the second device we developed, we created three evaluation exercise drills combining five basic movements, and conducted each of these twice at subjective high and low intensities with 10 university soccer players, collecting tracking data. In this basic performance evaluation, we confirm that these five actions and exercise intensity can be identified from the combined sensor data. In addition, we do not consider the first wristband-type device to be a target for evaluation because it has a simple design with a single accelerometer and the same measurement performance as the second wearable device. In order to instruct the players to perform relatively low-intensity exercises, we defined high intensity as 100% of their maximum effort, and low intensity as around 70%. The three exercise menus were created based on physical training and rehabilitation exercise menus for soccer, combining movements that often occur in field sports. The list of each exercise menu is shown in Figure . The players performed each movement in the order of the numbers in the figure, in the direction of the arrows, and used markers on the ground to turn and switch between movements. In order to check whether the movements recognized by the system using the pre-defined basic movement recognition rules were in the same order as the actual exercise menu, the recognition results were compared with the contents of each exercise menu and evaluated based on the matching rate. In addition, the basic movements were visually checked to see whether they were being performed according to the exercise menu, which was the true value. However, as there is variation in acceleration over a 5m moving range, the matching of the length of each movement was not evaluated.

4.3.3 Collecting Data for Evaluating Our Proposed Method

In order to evaluate the usefulness of the proposed method, we conducted an experiment to measure the movements of the players (excluding the goalkeeper) during a training match for a team of under-18 players that is part of a professional Japanese soccer team (J1 League) and analyzed the data using the proposed method. The subjects were 30 players aged between 15 and 18 years old, with 14 players in Team A, which was selected by the team's coach and had a high level of performance, and 16 players in Team B, which was made up of the other players. The matches were played three times, each lasting around 15 minutes, but actually lasted around 16 minutes, 18 minutes and 15 minutes respectively. There were 10 players on each team, excluding the goalkeeper, and substitutions were made between the three matches with the players not taking part so that all the players on the team took part. The match results were 2 - 0 in favor of Team A over the three matches. The time unit for calculating the moving entropy using the proposed method was set to one minute, which encompasses one play, based on the assumption that the team coach would evaluate the attacking scenes described below, and because the time that a team holds the ball during one attack in soccer is typically distributed widely over a range of 1 - 60 seconds [90]. The moving entropy between each player on the field was calculated for each minute, and a network diagram was generated.

4.4 Evaluation

4.4.1 Basic performance evaluation of prototype device

We evaluate the basic performance of the wearable device using basic movement identification. As it is clear that the five basic movements to be evaluated are related to the direction of movement on a plane relative to the body's orientation, as well as the speed and acceleration in the vertical direction, we designed the discrimination rules using the sensor data. The direction of movement on a plane relative to the body's orientation is calculated using the geomagnetic direction measured by the 9-axis sensor and the direction of movement measured by GPS, as shown in figure 4.6. The geomagnetic sensor can detect the direction of the N pole of the earth relative to the device's orientation, in the same way as a compass. In addition, GPS can detect the direction of movement based on the direction of the Earth's north pole. By superimposing these two sets of data based on the north pole, it is possible to calculate the direction of movement relative to the body if the device is fixed to the body. In addition, the speed of each movement was defined based on the general definition of movement, with forward running defined as 8 km/h or more, and side steps to the left and right and back steps defined as 4 km/h or more, which is about the speed of normal walking. The detection of jumps is based on the principle that the acceleration in mid-air is not affected by the gravitational acceleration of 1 G. A conscious jump is defined as a mid-air duration of 0.5 seconds or more. Based on the rules summarised in table 4.3, the average value of a 2-second window is applied, and the basic movement is determined by sliding it by 0.5 seconds.

4.4.2 Qualitative Comparison of Estimated Network Characteristics and Expert Evaluation

In order to examine the validity of the results of the proposed method, the match that had the most balanced attack and defense between the two teams (the third match) out of the three matches was selected, and the match was divided into three parts of about five minutes each. The match situation

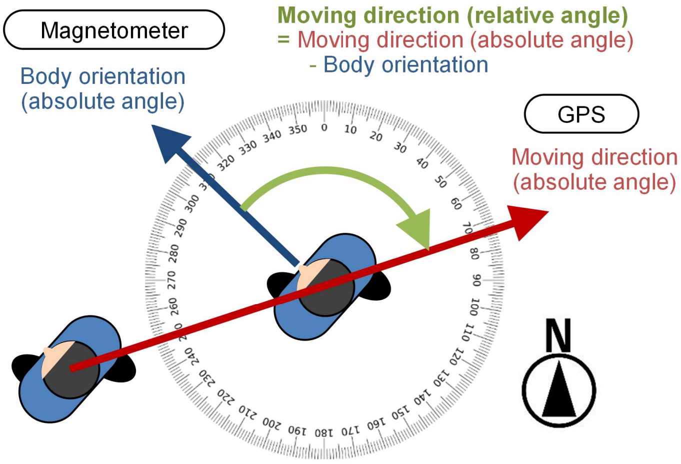


Figure 4.6: Calculation of relative angle of moving direction

Table 4.3: Rule definition of basic activity

Basic activities	Speed: v [km/h]	Moving direction: d [deg] (relative angle)	Acceleration: a [g]
Forward	v >8	0 <= d <= 60 or 300 <= d <360	-
Backward	v >4	120 <= d <= 240	-
Side (L)	v >4	240 <d <300	-
Side (R)	v >4	0 <d <120	-
Jump	-	-	Mean of a for 0.5 [sec] <0.5

was expressed simply, focusing on characteristic plays related to goal opportunities, and a qualitative comparison was made with the network diagrams. In this comparison, the authors presented the five corresponding network diagrams for each scene to the coaches, explained the characteristics of each network, such as the relationships between the players, and asked them to select the one that best represented the state of the game. The coaches and authors then worked together to confirm the validity of the selection.

4.4.3 Quantitative Comparison of Estimated Network Characteristics and Expert Evaluation

After the game, the team's coach visually checked the footage from the game, and scenes where the team was able to play in an organized manner that matched the team's coaching concept were extracted by the minute. The coaching concept of the target team was centered on moving the opposing defenders around by exchanging short passes to find points to attack, and the scenes that were extracted were all in line with this concept. There were scenes where there was a high possibility of scoring, such as scoring goals and taking shots, and 9 minutes were extracted from the total of 49 minutes. All of the scenes that were extracted were scenes where Team A held the ball and attacked. In comparing the results of the analysis, we calculated the average degree for each team per player, the average degree for connections within the team only, the average degree for connections from team A to team B only, and the average degree for connections from team B to team A only, using the transfer entropy for each minute of the network, and then evaluated whether there was a statistical difference between the extracted scenes and the other scenes for each of these indicators, and considered the interpretation of the results. To evaluate significant differences, we used Welch's t-test as a non-paired two-group t-test, with a significance level of 0.05.

4.5 Result

4.5.1 Basic performance evaluation of prototype device

We evaluated the results of basic movement identification based on sensor data from wearable devices in three drills. In the results of the exercise drill (A), we were able to confirm that the direction of movement of the players, as measured by GPS, changed to the correct direction for each movement, and that the speed of movement exceeded the threshold set by the rules. In the results of the exercise drill (B), we were able to confirm that the direction of the body, as measured by the geomagnetic field of the 9-axis sensor, changed by 180 degrees when the body was turned over, but that the direction of movement did not change, and that the direction of the body did not change when only the direction of movement was reversed. The results of the exercise drill (C) showed that the changes in direction and speed were reasonable, and that the acceleration during the jump decreased to nearly 0 G, as set in the rules. The forward running had an accuracy rate of 95.8%, while the backward running, left side step and right side step had 100%, and the jump had 92.5%. However, in the low-intensity exercise drill (B), there were some parts where it was clearly not possible to run, so these were excluded from the evaluation. In addition, the reason why forward running could not be detected was that when entering the body inversion movement, the distance of the forward running before and after it became shorter, and the speed did not increase to the threshold value. In addition, in the case of jumps, the reason was that the duration of the jump was shorter than the set 0.5 seconds for some players. Apart from the parts where the movements we were assuming to be discriminated against were not performed for these reasons, the system was able to discriminate correctly, and since it can be considered that if the basic movements are performed with the travel distance and duration of flight set in this evaluation, it can be considered that it can identify 100%, it was confirmed that it has sufficient performance for field sports measurement such as soccer.

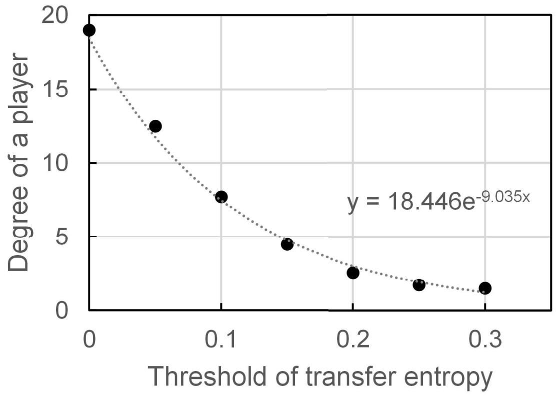


Figure 4.7: Degree of a player by setting threshold (P) of transfer entropy

4.5.2 Qualitative Evaluation of Cooperation Graphs

In creating a network diagram based on the transfer entropy between players from the measured data, we set the threshold value P to generate the effective graph shown in figure 4.7. To set the threshold, we calculated the moving entropy between all the players who were competing each minute, and then calculated the average degree per player when the threshold P was changed between 0 and 0.3, as shown in figure 4.7. As P increases, the degree per player decreases exponentially. In this study, as described in the previous chapter, the degree was set to P=0.1 in order to keep it below half of the total connection. This threshold was applied to the creation of all network diagrams.

We confirmed that the network diagrams created using the proposed method expressed the state of the match appropriately. The commentary on the state of the match in the early stages of the game was as follows 'Team A mainly kept the ball and continued to pass it around. Team B, especially numbers 34 and 36, aggressively pressed from the front line, but they conceded a goal after being attacked from the open space on the side that was created from the pass by Team A's number 10.' The corresponding time network is shown in figure 4.8. Looking at figure 4.8, we can see that B Team No. 34 has the highest degree of 6 between B Team and A Team, so we can say that he is moving and defending ahead of the other players in B Team. We can also see that there are more arrows within A Team (average degree within A Team: 4.4, average degree within B Team: 2.9), so we can see that there was a lot of information exchange within the team. Furthermore, as the arrows heading towards No. 10 on Team A are concentrated, it has the highest degree, at 10. From this, it is not possible to determine whether No. 10 on Team A was passing the ball a lot, but it can be seen that by moving in line with the players pressing and the players around them, they made it easier to receive the ball and were actively involved in the team's play.

The following is the coach's commentary on the midfield scene: 'There was a period of intense attacking and defending on the left side of Team A and the right side of Team B. There were many

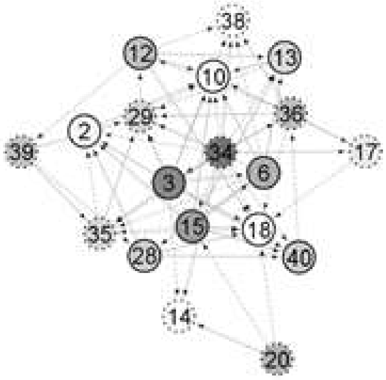


Figure 4.8: Network showing MF (34, 36) in team B pressing players in team A

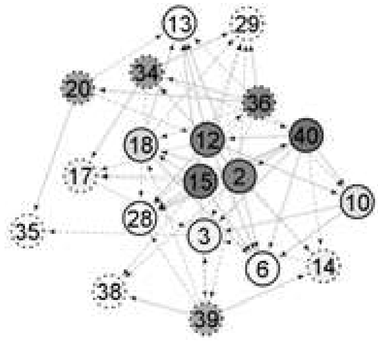


Figure 4.9: Network showing defense of MF (38) in team B responding to offence of MF (15) in team A on the side of the court

scenes where the defenders responded to counter-attacks, with passes being played into the space in front of No. 15 on the left side of Team A, and Team B's defenders (DF) No. 38 returning to defend. The network during the time of response is shown in figure 4.9. The attacking player on the left side of Team A was No. 15, and the corresponding players on the right side of Team B were Nos. 38 and 39. In figure 4.9, the arrows connect the players on the same side (No. 15 to No. 38, No. 15 to No. 39), showing that the players on the same side of Team B (No. 38 and No. 39) were moving in response to the movements of Team A's No. 15.

Finally, the instructor's commentary on the final scene was as follows 'Team B maintained and controlled the defensive line in a high position, while Team A mainly held the ball in a low position and continued to look for openings to attack. There were many scenes where Team A's forward (FW) No. 40 looked for the right moment to receive a pass while moving in line with Team B's defenders (No. 20 and No. 29), and there were also many scenes where both sides of Team A looked to attack from behind.' The corresponding time period network is shown in figure 4.10. In figure 4.10, it can be seen that there are many networks from the B team players to the A team players (average degree

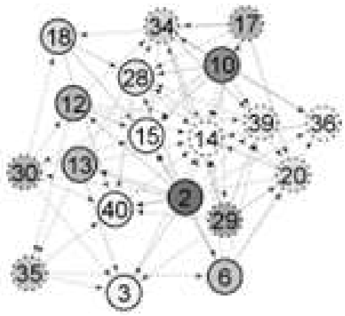


Figure 4.10: Network showing FW (15) in team A moving in response to DF (20, 29) in team B

of B to A: 4.9, average degree of A to B: 4.4, average degree of A to A: 3.2, average degree of B to B: 2.7). In particular, the arrows connect from the DF (No. 20 and No. 29) of Team B to the FW (No. 40) of Team A, showing that the DF of Team B moved first to adjust their position, and the players of Team A moved while watching that movement. As described above, by comparing the subjective evaluation of the situation by the coach with the network diagram of the same time, we were able to confirm that the relationships between the main players involved in the situation were reflected in the network.

4.5.3 Quantitative Evaluation of Cooperation Graphs

In order to confirm whether it is possible to actually represent the play that the coach intends by quantifying the organizational performance using the proposed method, we evaluated whether there were any statistical differences in the degree, which is a typical network analysis index, between the scenes (total 9 minutes) extracted by the coach from the video of the practice match and other scenes. In the box plots showing the distribution of the data for each evaluation, the lower limit of the minimum value was set at the first quartile minus 1.5 times the interquartile range, and the upper limit of the maximum value was set at the third quartile plus 1.5 times the interquartile range. First, to check for differences in the characteristics of the two teams, etc., we compared the average degree of each player in both teams in all scenes, as shown in figure 4.11. There was no significant difference between the two teams in terms of the average degree of all the players in both teams, the degree of the players who received arrows, and the degree of the players who sent arrows, and it was confirmed that they were equivalent. In addition, in figure 4.12, the degree of the network within each team is compared, and there was no significant difference and they were equivalent. From this, Team A won, but in terms of the characteristics of the information transmission network, it is not possible to see any difference in the average values throughout the match. In addition, given that the players from both teams trained together and received the same team policy guidance, it can be thought that the characteristics of their play were relatively similar.

Next, we compared the differences between the scenes where the coach judged that the team was

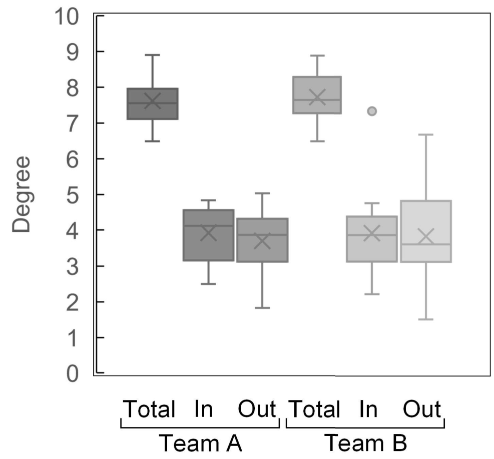


Figure 4.11: Comparison of average degree (total network) of each player between both teams

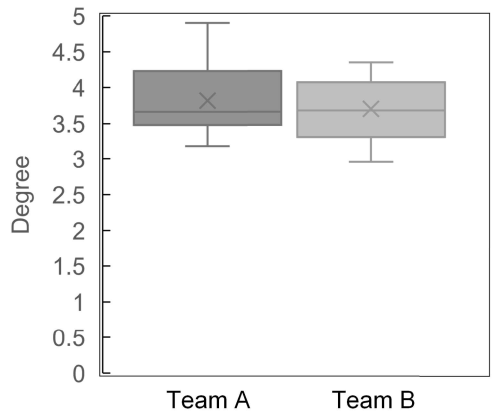


Figure 4.12: Comparison of average degree (team network) of each player between both teams

able to carry out an organized attack and the other scenes. Figure 4.13 shows the average degree of the overall network of the players from Team A in each scene, as well as the average degree of the intra-team network of the players from Team A. There was no significant difference between the two scenes in either of these measures. In other words, even in the scene where it was assumed that an organized attack was being carried out, there was no difference in the overall network of information transmission or the indicators within the team. In order to evaluate the difference in information transmission between the teams, figure 4.14 compares the average degree of each player from Team A to Team B and from Team B to Team A in both scenes. The comparison showed that the degree of information flow from Team B to Team A was significantly higher in the scenes where the team was able to carry out an organized attack. This suggests that the degree of information flow from the opposing team to the friendly team can be used as an indicator to evaluate scenes where the target

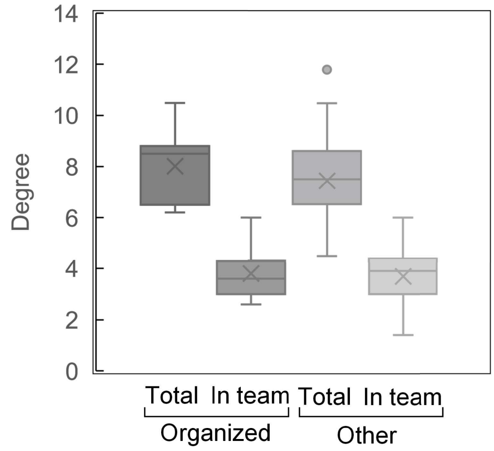


Figure 4.13: Comparison of average degree of each scene between organized scenes and the other scenes in team A

team is able to carry out an organized attack.

4.6 Discussion

4.6.1 Validity and usefulness of the proposed method

In this study, we evaluated whether it is possible to visualize the organizational performance of players and teams by analyzing the information transfer between players using the proposed transfer entropy. In the qualitative evaluation, we compared the situation of the game as seen by the coach with the network diagram, which is the analysis result, and confirmed that the structure of the network and the connection of the graph between players adequately expresses the relationship between the players who are mainly involved in the situation of the game. When evaluating team performance, the number of successful passes and shots are generally used in soccer. Research has been conducted to analyze the direction and length of passes from video footage in order to evaluate passes in detail [91]. In ball games, passing is the most important team play, but as only one player holds the ball in a team sport, there are more players other than the passer and receiver, and ‘off-the-ball movement’, which is the preparatory movement that creates a pass, is considered important [92]. In addition, it is said that each player checks the positional relationship with the opposing player and plays a waiting game, timing the movement to receive the pass. If the information transfer shown in the proposed method of this study is interpreted to reflect off-the-ball movement, it may be possible to use it in the future as an indicator of the liveliness of off-the-ball movement in the coaching and tactical planning of team sports. On the other hand, the current method does not measure the positions or fields of view of the players, and there is a limitation in that it is not possible to directly determine whether there was any interaction between the players. In the future, if it becomes possible to measure the positions

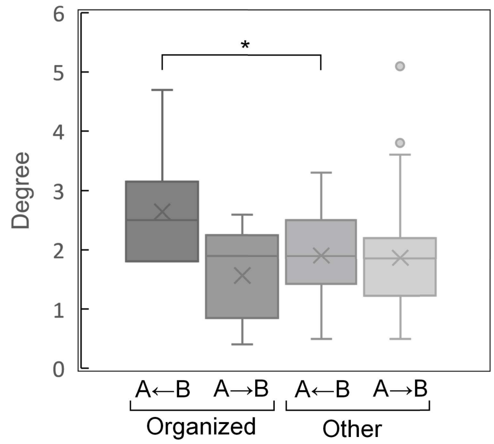


Figure 4.14: Comparison of average degree from a team to the other team of each scene between organized scenes and the other senses

and body orientations of the players using GPS or geomagnetic sensors, it will be possible to create visualizations that take into account the degree of direct influence. In the quantitative evaluation of network indicators, it was confirmed that, in the systematic attacks intended by the target team (Team A), although there were no changes in overall order, etc., there was a statistically significant increase in information transmission from the opposing team (Team B). The tactics intended by the target team were for the attacking team to pass the ball around as much as possible so that the opposing team would not get the ball, and to look for gaps in the defense. In other words, it seems that the players receiving the pass create opportunities by moving in conjunction with the positions and movements of the opposing team. This is consistent with the results of the analysis, which showed that there was a lot of information transfer from the opposing team. In previous research on motion analysis in soccer, it has been said that the players on the defensive team react to the movements and passing routes of the attacking team and steal the ball, and that there is an increase in the number of movements that react specifically to the stealing of the ball [93]. Therefore, in order for the defenders to be able to win the ball more often, it is thought that the result would be that there would be more information being communicated from the attacking side. In this respect, it can be seen that Team B, which lost after a long period of defending, was not able to effectively make moves to win the ball. In the practice match we analyzed, Team A held the ball for a long time, so it is easy to understand the analysis results as Team A attacking and Team B defending. In order to evaluate matches in which the offense and defense are more evenly matched and there is a lot of switching between the two, it will be necessary to consider other network indicators, as well as dividing up scenes according to the ball possession rate and the positions of the players, and making the time divisions variable. In addition, the transfer entropy of the proposed method only considers the impact on the immediate time (fixed-width delay) in the relationship between the movements of two players, but we also think that it is necessary to extend it to transfer entropy that includes the impact of more delayed time [94, 37].

4.6.2 Practical Application of the Proposed Method

The proposed method in this study analyses the relationship between independent acceleration data, and it is thought that the computational processing of the analysis can be easily implemented in existing measurement systems. Currently, the most widely used devices in professional sports such as soccer and rugby are those that analyze data collected wirelessly in real time or data collected collectively via a wired connection on a cloud or PC, and then report the results for coaches and players to check. The data used in the proposed method is acceleration intensity data collected at one-second intervals, so it is a volume of data that can be achieved using either wireless or wired communication and it can be implemented simply by adding calculation and display functions to the cloud, so it can be said to have a high degree of affinity. In addition, the method of visualization can be made easier for the coach to understand by reflecting the actual positions of the players obtained using GPS and other methods in the network diagram, and by combining existing indicators such as distance moved. The size and color of the circles used to represent the players do not have to be limited to the two indicators shown in figure 4.3, and for example, for the purpose of evaluating the players who are important as routes connecting the network, it is also possible to use measures such as mediation centrality. When the proposed method is systemized, it is possible to provide a user-friendly interface that can be used for a variety of purposes by selecting the desired indicators from a list of candidates on the display screen of a PC or tablet, combining various expression methods such as color and node size, and drawing the indicator values on a graph. The system that was actually developed is used to visualise the degree of coordination between children and between children and teachers during physical education classes, and it is used to evaluate the degree to which the objectives of the class have been achieved and the proficiency of the teacher [95]. In the same way as the qualitative evaluation, the practical use of the network diagram visualizing information transmission can be used to evaluate the scenes extracted based on the knowledge of the instructor in a more quantitative and objective way. Furthermore, it is thought that the system can be used to extract similar scenes that the instructor has missed by collecting network information on the scenes extracted by the instructor and having the system learn the characteristics. In this study, we used a wristband-type device to measure the movements of the athletes, but we can assume that similar results can be obtained using a device that is attached to the torso on the back, as both devices measure values that correlate with the intensity of the athlete's full-body movements, as is the case in professional sports. Wristband-type devices are used to measure the exercise of a wide range of athletes [96], and they can also be used to manage conditions by measuring sleep and activity levels in everyday life. They are also low-cost and easy to wear, so they are considered to be suitable for use in education and sports, and for athletes in their formative years. On the other hand, because the use of wristband-type devices is not currently permitted under competition rules, there is a restriction that they cannot be used in official matches, so in this study, the sensors were protected by a cloth band, and after confirming with the coaches and athletes that there were no safety issues, they were used. For official matches that are not covered by the wristband device, we are considering measuring in accordance with the rules of the game by attaching the sensor to the back of the player using a vest with pockets. In the future, it is thought that similar analysis can be achieved through non-contact measurement, not just with wearable devices, but also by utilizing the recognition of each player's position and movement through video analysis.

4.7 Conclusion

The following conclusions were reached as a result of the experiments and evaluation of the analysis index for latent information transmission using transfer entropy proposed in this study. Transfer entropy was calculated from the acceleration data of the wearable devices worn by each player on the soccer team, and the structure of the network was visualized. It was confirmed that the structure of the network visualized in the multiple scenes extracted was consistent with the commentary of the team's coach and was valid. Furthermore, in the scenes where the team leader had intended to carry out an organized attack, it was confirmed that the information transmission from the opposing team was significantly greater, suggesting that this could be used as an indicator of the organized performance of a team sport. As a result, it is expected that the introduction of the proposed method will make it easier for on-site leaders to quantitatively and objectively understand the relationships between the movements of the players during training . By collecting long-term data, we hope to develop a system that can be used to understand changes in training effectiveness and player development based on data in the future. In addition, the proposed method is easy to implement in wearable devices and analysis systems that are popular in professional sports, etc., and we will continue to work on tuning the calculation logic, expanding the network indicators, and supporting the meaning of each indicator with the aim of using it in more teams in the future.

Chapter 5

Basketball Score Prediction from Trajectories using GNNs Based on Player Cooperation Graphs

5.1 Introduction

Analysis of location information acquired in daily life enables a detailed understanding of human behavior and contributes widely to society. Trajectory data, accumulated in large volumes, has attracted attention from many fields, including computer science, sociology, and geography, and analysis methods have been actively researched. Various data processing methods have been developed and applied to understand trajectory data. In order to understand and utilize trajectory data, various data processing methods have been developed and applied. In recent years, deep learning, a non-linear machine learning method that does not require feature extraction based on human knowledge, has become mainstream. Deep learning methods used for trajectory data analysis include RNN (Recurrent Neural Network) [54], which learns time series information, and a method that combines RNN and CNN (Convolutional Neural Networks), which takes image data as input [56]. To understand behavior through the analysis of trajectory data, new research has focused on extracting meaning from multiple trajectories based not only on the behavior of individuals but also on data from crowds of unspecified people [57]. In addition, the team sports that are the focus of this study are the most suitable subject for research on understanding group behavior. A spatiotemporal analysis method for group trajectory data has been proposed for scene evaluation in team sports [10]. Methods using deep learning have been studied, such as predicting scores [13] using deep learning of RNNs and CNNs. However, while previous research has been able to predict performance such as scoring opportunities from the trajectory of the group, it has not been possible to quantitatively evaluate the degree of influence of individuals on performance while taking into account the interaction between players in the group. Previously, we proposed a method for evaluating individual contributions in basketball using a deep learning model that predicts scoring opportunities using images of the trajectories of the movement of the positions of the players and the ball as input [97]. This method was a deep learning model that was trained using a fully connected network, with the trajectory images of each player input into a different CNN, as ablation was applied to the input data. As a result, it was not possible to consider the interaction between players or the positional relationship between players and the ball, and it was not a model that could realistically represent team sports. In this paper, we propose a method to obtain higher prediction accuracy and a better understanding of the relationship between organizations by learning the mutual influence of players and balls in a deep learning model with player and ball trajectory data as input. The proposed method introduces GCN (Graph Convolutional Networks), a type of GNN (Graph Neural Networks), to conventional DNNs such as CNNs and RNNs, which learn the time series of trajectory data by providing a graph structure of the relationships among multiple input data. GCNs are convolutional neural networks that take graph structures as input and are used in various fields such as text classification [98] and compound classification [99]. In the study of the proposed method, we apply the NBA professional basketball tracking data, which is publicly available, as a large amount of data is required for training deep learning models. As a result, we report that introducing GCN improves the prediction accuracy of the scoring opportunity prediction model trained on basketball tracking data, and leads to a better understanding of the contribution of each player in team sports. We also discuss the impact of graph structure on the prediction accuracy of GCN and future applications of this method.

5.2 Method

We designed a neural network to generate a prediction model by learning a dataset of player and ball trajectories tagged with scoring/not-scoring for each attacking scene. In addition to RNNs and CNNs that use time-series data as input, we applied GCNs, which can define the relationships among input data as a graph structure. Unlike the convolution process for image pixels, a convolution process with a graph structure in which each node has a different number of neighboring nodes has been proposed [19]. Figure 5.1 shows an overview of the convolution process. In a graph structure with N nodes, the convolution process for a node feature $\boldsymbol{f}_{in} \in \mathbb{R}^{N \times C}$, where the feature of each node is C-dimensional, is as follows. The output $\boldsymbol{f}_{in} \in \mathbb{R}^{N \times F}$ of the number of feature maps (F) is calculated by the following formula (4) using the adjacency matrix $\hat{\boldsymbol{A}} \in \mathbb{R}^{N \times N}$ and the weight matrix $\boldsymbol{W} \in \mathbb{R}^{C \times F}$, which are available from the graph structure.

$$f_{out} = \boldsymbol{\Lambda}^{-1/2} \hat{\boldsymbol{A}} \boldsymbol{\Lambda}^{-1/2} f_{in} \boldsymbol{W} \tag{5.1}$$

$\hat{bmA}$is the sum of the adjacency matrix $\boldsymbol{A}$ and the identity matrix, which is the loop of the graph structure. The adjacency matrix $\boldsymbol{A}$ is a matrix of node connectivity, where each element $\hat{A}_{ij}$ is 1 if

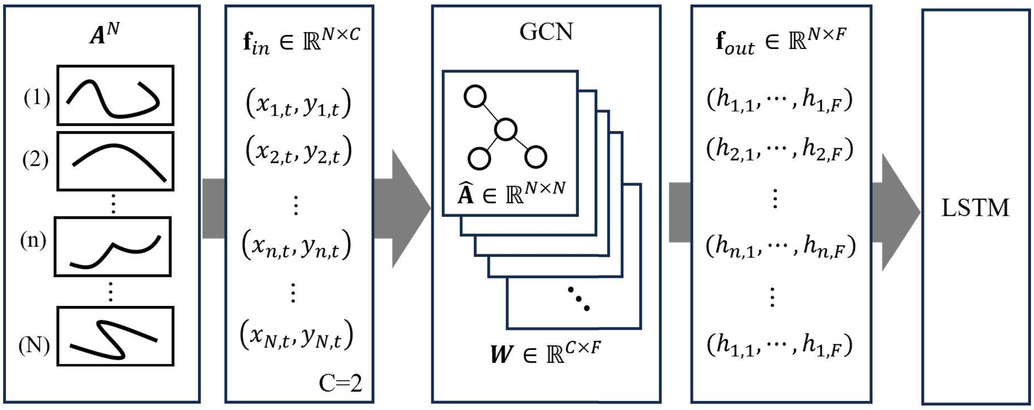


Figure 5.1: Overview of graph convolutional network

node i and node j are connected, and 0 otherwise. $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal components are the eigenvalues of the graph laplacian, expressed in terms of the order of the nodes.

We designed a neural network structure applying GCN, which is shown in figure 5.2. This neural network learns the presence or absence of scores in the scene of the input trajectory data. Supervised data is 1/0 indicating tag scoring/not scoring, and the model is trained as a binary classification model. In the designed network structure, the trajectory data of the player and the ball are input as time series data in 2D coordinates, and after convolution processing based on the graph structure by GCN, the output is input to LSTM, a type of RNN that can learn the features of time series. The output of the LSTM is input to its fully connected layer, and the final output is the probability of scoring a goal by softmax. Each player and ball input is one node of the graph. The structure also allows each input to be ablated when calculating the contribution of each player by our proposed method. The parameters of each layer of the deep learning model were determined based on prior tuning and previous research. The length of the input data was set to 100 to match the minimum number of frames for each attacking scene to be prepared as a data set. The training phase was performed for a maximum of 100 epochs and stopped when the learning loss stopped decreasing. The batch size was 64, which was adjusted to be the fastest learning speed and the value at which learning converged. During training, 1/10 of the training data was used for validation.

Figure 5.3 and 5.4 show the generation of time-series data for the input movement trajectory. The long and short sides of the basketball court are represented by x and y, respectively, and normalized so as to have a minimum value of 0 and a maximum value of 1. The coordinate data is rotated and transformed for an attacking direction with a larger x value. The position of a player or ball is represented by a vector (x, y). The positions of players and balls are represented as (x, y) vectors, i.e., arrays of vectors representing positions for sequential frames in the time direction, and are input to the deep learning model as time series data in 2D coordinates.

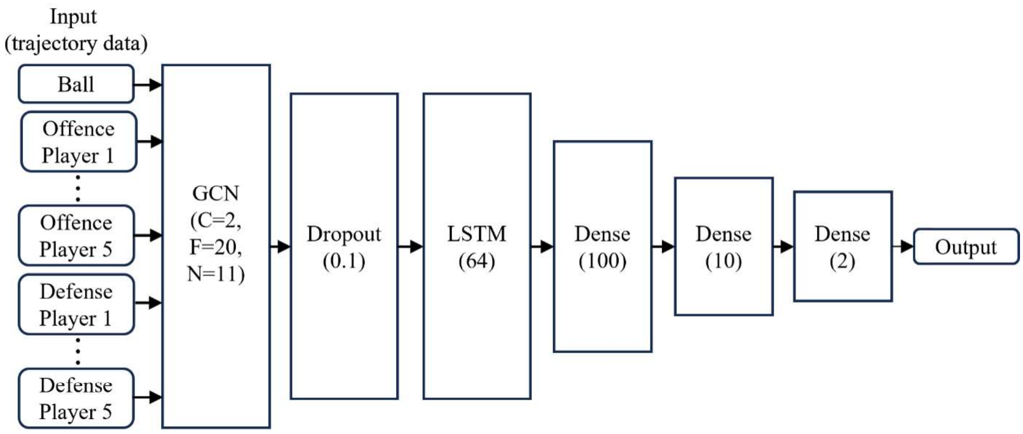


Figure 5.2: Design of neural networks

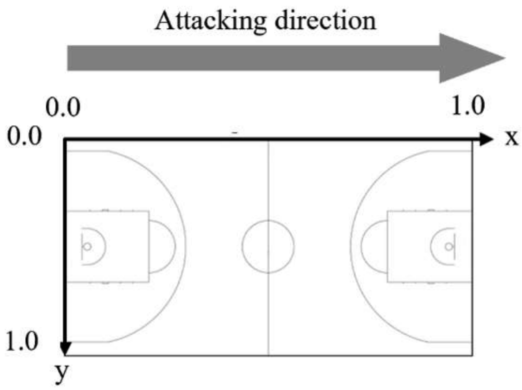


Figure 5.3: Coordinate system of player and ball positions on a basketball court

5.3 Dataset

In this paper, we use the official tracking data of the NBA, the USA's professional basketball league, from the 2015-2016 season, which has been used in various studies on basketball. The data is recorded by the SportsVU camera system, which extracts the trajectories of players and a ball by image recognition from the videos captured by multiple cameras. The data frame rate is 25 fps, consisting of 2D positions of 10 players and a ball covering an area of 94 feet * 50 feet of the basketball court for 663 regular season games, excluding playoffs and other events. Each player's and ball data is distinguished by a unique ID and can be linked to event data such as shots and fouls. The event data is recorded with the tracking data and includes events such as fouls, ball outs, and steals that can be used to determine scores and turnovers with timestamps. In this paper, we use 102,452 attack scenes, of which

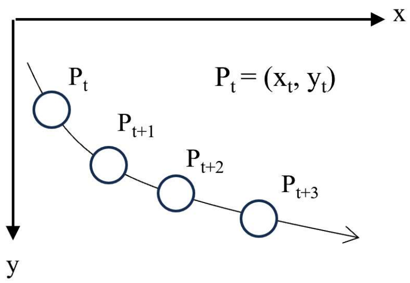


Figure 5.4: Representation of time series data of player and ball position coordinates

51,537 succeeded in scoring, as shown in figure 5.1.

Table 5.1: Number of data in the basketball dataset

	Label		
	Scoring (1)	Other (0)	Total
Training	41,275	40,686	81,961
Test	10,262	10,229	20,491
Total	51,537	50,915	102,452

The NBA tracking data applied in the study was converted into a dataset necessary for the training and evaluation of the proposed method. The positional coordinates are time-series data recorded at a sampling rate of 25 Hz, with the number and positional coordinates of the participating players recorded for each sample, as well as the positional coordinates of the ball. In basketball, players can be replaced as many times as necessary during breaks in play, so the players are frequently replaced. The event data contains the time when the event data occurred and the number of players involved. To train a model that predicts the likelihood of scoring opportunities for the proposed method, the data is divided into data units by the occurrence time of turnover events, such as scoring by shooting, steals, fouls, and other offensive and defensive turnovers, as shown in figure 5.5, to determine whether the turnover was caused by scoring or other events. The data were tagged with a binary value of scoring (1)/not-scoring (0). Players are replaced during the game, but not except in events where turnovers occur, so that the same five players from both teams are playing in the separated segments. In the analysis in this paper, a dataset was generated from the target public tracking data (663 games in one season), tagged with scores for 100 frames (about 4 seconds) or more of what is considered to be a valid play. However, in this model, for data longer than 100 frames, only the 100 frames with the slowest time are used as input data, and the lengths of the input data are aligned. Since the positional relationship between the player or ball trajectory and the basketball court is always constant, information on the positional relationship with the goal is also included. Therefore, in attacking scenes where the offensive

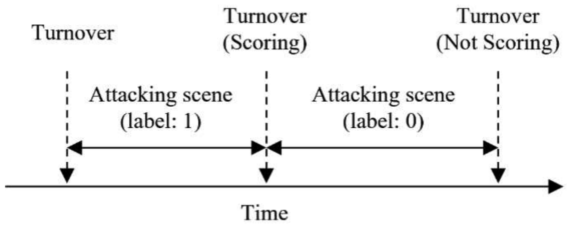


Figure 5.5: Time delimitation and labeling of attack scenes

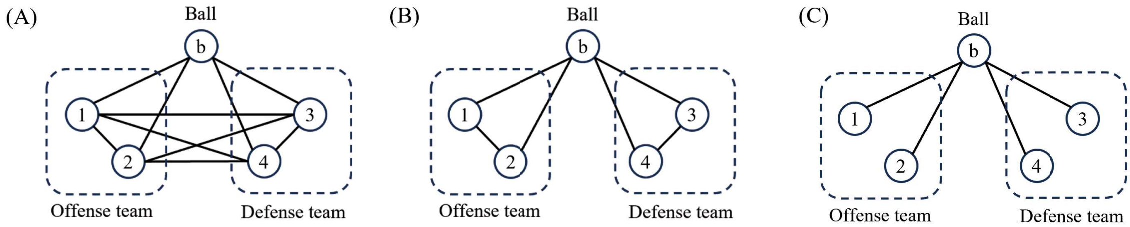


Figure 5.6: Graph structures of the connection between the ball and players (Examples with 2 players on one team)

team's players or a ball is far from the goal, there are few chances to score, making prediction easier. On the other hand, when players or a ball is close to the goal, the possibility of scoring cannot be simply predicted.

5.4 Evaluation and Result

To improve the accuracy of the proposed method, the evaluation of players' contribution to scoring opportunities, this paper used training data extracted from NBA tracking data to train a GCN model to predict the probability of scoring opportunities and calculated the prediction accuracy using test data. For this evaluation, we assumed the two graph structures shown in figure 5.6. We evaluated the prediction accuracy of the deep learning model for each graph structure, assuming two graph structures: full-linkage, in which all players and the ball are connected, and intra-team linkage, in which players within a team and the ball are connected. By evaluating the difference in prediction accuracy between the two graph structures, we believe that the graph structure with higher prediction accuracy is closer to the interaction of player positions in actual basketball. We also confirm that the introduction of the graph structure improves the prediction accuracy compared to the previously reported CNN-based model and that it is applicable to the proposed method.

To evaluate the feasibility and prediction accuracy of the proposed deep learning model using GCNs, we input a test dataset and compared the prediction results with the results actually scored by the model. As accuracy metrics, we calculated the accuracy, recall, and F-measure for each label, calculated the mean (Macro), and added it to the AUC, which compares model accuracy, and summarized the results for three different graph structures in Table 5.2, Table 5.3 and Table 5.4. As shown in the table, the graph with the full-linkadge network had better prediction accuracy, with an AUC above 0.95. In other words, it was confirmed that considering the interaction between the attacking and defending players also contributed more to the accuracy. We also confirmed that this accuracy was improved compared to the reported CNN-based model. In other words, it was confirmed that the application of the GCN, considering the relationship between the player and the ball, contributed to the improvement of the prediction performance of scoring chances.

Table 5.2: Prediction accuracy values of scoring opportunities for graph structure (A)

Label	Precision	Recall	AUC
1	0.879	0.906	0.951
0	0.893	0.885	

Table 5.3: Prediction accuracy values of scoring opportunities for graph structure (B)

Label	Precision	Recall	AUC
1	0.876	0.813	0.918
0	0.810	0.872	

Table 5.4: Prediction accuracy values of scoring opportunities for graph structure (C)

Label	Precision	Recall	AUC
1	0.830	0.892	0.911
0	0.870	0.800	

5.5 Discussion

Toward the evaluation of player contribution using the model proposed in this study for predicting scoring opportunities in basketball, we have achieved that the GCN model, which takes into account the interaction between players, can predict scoring opportunities with higher accuracy than the previously reported CNN-based model [16]. Furthermore, the high prediction accuracy obtained with GCN suggests that the mutual influence of the positional relationship between players and the ball in team sports may have an impact on scoring opportunities. In particular, the accuracy of the model using the all-combination graph was higher than that of the model using intra-team combinations, suggesting that the influence of players from different teams contributed to the higher prediction accuracy. In the future, it is possible that new measures of organizational performance could be obtained by evaluating changes in accuracy due to various graph structures. For example, if it were possible to quantify the impact of missing graph structure among some players, it would be possible to evaluate the impact of coordination among players on organizational performance. On the other hand, several challenges remain in the practical application of deep learning models for predicting the probability of scoring opportunities. In the dataset used in this paper, only the last 100 frames of the segmented scenes were used as input data, and there were many data with clear differences between the trajectories of scenes in which no goals were scored and those in which goals were scored, and the two could be discriminated with some ease. Future work should consider the use of data sets that focus on various scoring processes, such as variable-length data input and the use of trajectories in the middle of each playing scene. Our proposed method can also replace the human subjective evaluation of team and player performance with data-only evaluation, which is both objective and greatly improves efficiency. If it becomes possible to visualize the linkages, it will be possible to provide content that will attract viewers' interest and enable a deeper understanding of sports play in various professional sports broadcasts in which existing tracking systems are in operation. If it becomes possible to extract and visualize the connections between players who are active on the field where scoring chances are likely to occur, it will be possible to provide viewers with content that will attract their interest and give them a deeper understanding of sports play. In the future, models and analytical methods learned in professional sports can be used to develop the next generation of athletes, which is expected to dramatically improve the efficiency and effectiveness of their training.

5.6 Conclusion

In order to improve the accuracy of research using deep learning models that predict scoring opportunities based on the trajectory data of basketball players and balls as input, we proposed the application of GCN, which can express the relationship between players and balls in graph structure. We used trajectory data for each play situation extracted from tracking data for one season (663 games) of the NBA, a professional basketball league, as input, and trained a deep learning model on 80% of the data, and evaluated its accuracy on the remaining 20% of the data. The proposed method achieved a high accuracy of AUC=0.95 compared to a model that does not consider the interaction between players. From these results, it can be expected that the model learned using the proposed method more accurately reflects the behaviour of players in team sports, and that it will also be more accurate in evaluating the contribution of players. It was also suggested that setting the graph structure that expresses the relationship between players and the ball more correctly would contribute to improving accuracy.

Chapter 6

Conclusion

In this study, we introduced a method for evaluating player performance and interactions using AI technology. This approach enables the prediction of team performance and other aspects by employing a deep learning model that considers player cooperation based on positional information and movement data. The model is subsequently interpreted to provide meaningful insights.

We proposed three main methodologies. First, we proposed a deep learning model that predicts team performance based on the trajectories of players and the ball in sports such as basketball and soccer, and a method for obtaining explainability by adding perturbations to the inputs of this black box model. Second, we proposed a method for evaluating the interaction between players based on information theory that examines the relationship between the movements of two players. Third, we proposed an extension of the first method using a neural network that can learn the mutual positional relationship between players.

We tested these methods using tracking data from professional basketball and football league matches. For the first method, we developed a deep learning model using a convolutional neural network (CNN) to estimate the number of points scored from the trajectories of players and the ball. This resulted in a high-precision scoring prediction model for basketball, achieving an area under the curve (AUC) of approximately 0.92. The contribution of each player was evaluated by assessing the change in output values when the input of each player was masked, and the validity of the results was confirmed by comparing the average values for one season with existing evaluation metrics. In particular, it was found that the correlation with the efficiency, a measure of the contribution of basketball players, was significantly high, while the correlation with other individual scores such as points and assists was low. A model was also constructed for predicting scoring opportunities in soccer, and a prediction accuracy of approximately 0.87 was achieved, although the dataset was smaller than that for basketball. In addition, when evaluating the contribution of players to scoring opportunities in soccer, we evaluated the correlation with individual scores for each team, and found that it was highly correlated with the number of offensive plays that had a large impact on the movement trajectory, such as shooting, receiving passes, and dribbling. From these results, we found that the proposed player contribution evaluation is a new method that evaluates the direct contribution to scoring using explainable AI, and that it is also possible to interpret the validity from the conventionally used indicators.

In the second method, we used the concept of transfer entropy from information theory to calculate the degree of cooperation between players in a soccer match. To calculate this, we proposed a method of converting time series data such as the independent acceleration and movement of two people into continuous probability variables, and evaluating whether there is cooperation between the two people. In addition, in order to collect evaluation data in actual team sports, we designed and prototyped a wearable device that can collect multiple sensor data with little burden on the players. Using the data collected from the football matches using the prototype device, we compared the network structure based on transfer entropy with the organised performance of the matches as evaluated by experts. As a result, we were able to confirm that cooperation between teams in particular has a significant impact on organised performance.

In the third method, we applied a Graph Neural Network (GNN) to the deep learning model of the first method based on the hypothesis identified in the second method. By incorporating interactions between players, we aimed to improve the accuracy of the basketball scoring prediction model and improve the validity of the player contribution evaluation with a more realistic model. Using a basketball dataset, we evaluated the prediction accuracy using three different assumptions about the graph structure between players. As a result, we confirmed that the highest-precision scoring prediction model, with an AUC of 0.95 or higher, was obtained by considering the interaction between all players and the ball. In addition, we confirmed that the prediction accuracy was higher than that of the model using CNN that did not consider the interaction between players, and that the precision and recall were also close to the level required in sports, at around 0.9.

The results of this study show that AI models can effectively evaluate direct contributions such as scoring in team sports. By incorporating interactions between players into the model, the accuracy of team performance prediction has improved, and it has become possible to evaluate individual contributions more accurately. By using AI models, it is possible to automate evaluations that were previously carried out empirically by coaches and experts, and it is also possible to develop new tactics and re-evaluate players who have been overlooked.

However, this research is limited by the input data and model construction process. We relied solely on player trajectory data, which, while useful for extracting sports performance from positional information, does not account for ball skills or physical abilities not reflected in positional data. Interviews with players have shown that even when occupying the same position, situational judgment varies depending on physical abilities and skills. Future AI model developments should aim to incorporate these additional inputs, creating accurate predictive models of player behavior in complex team sports.

As AI research progresses, we anticipate a clearer understanding of the complex spatial recognition, decision-making processes, and mechanisms that allow athletes to demonstrate their physical abilities in crucial moments. This will not only enhance evaluations and understanding of team and athlete performance but also predict developmental processes and lead to more efficient training methods. Future sports research outcomes will contribute to children's education and lifelong sports well-being, highlighting the potential of AI-driven data analysis technology in the sports field.

Chapter 7

Acknowledgment

I would like to express my deepest gratitude to Professor Hirozumi Yamaguchi and Associate Professor Akira Uchiyama from the Graduate School of Information Science and Technology, Osaka University, for their invaluable guidance and support throughout my doctoral research. Their expertise and mentorship have been instrumental in the completion of this dissertation.

I am also sincerely grateful to Professor Takuya Magome from Otemon University for his insightful advice based on his extensive experience in the field of sports. His suggestions significantly enriched the practical aspects of this research.

In addition, I am sincerely thankful to the staff of Kashiwa Reysol's youth team for their cooperation during the experiments. Their willingness to assist and provide access to valuable resources was essential for the empirical aspects of this study.

Finally, I would like to acknowledge the members and staff of the Yamaguchi Laboratory at Osaka University for their constant support and collaboration throughout the research period. I am also indebted to the members of the Research & Development Group at Hitachi, Ltd. for their assistance and encouragement.

Thank you all for your invaluable contributions and encouragement.

Reference

[1] Andrey Rudenko et al. "Human motion trajectory prediction: A survey". In: *The International Journal of Robotics Research* 39.8 (2020), pp. 895–935.

[2] Huansheng Song et al. "Vision-based vehicle detection and counting system using deep learning in highway scenes". In: *European Transport Research Review* 11.1 (2019), pp. 1–16.

[3] Wang-Chien Lee and John Krumm. "Trajectory preprocessing". In: *Computing with spatial trajectories*. Springer, 2011, pp. 3–33.

[4] Zuchao Wang et al. "Visual traffic jam analysis based on trajectory data". In: *IEEE transactions on visualization and computer graphics* 19.12 (2013), pp. 2159–2168.

[5] Claudio Piciarelli, Christian Micheloni, and Gian Luca Foresti. "Trajectory-based anomalous event detection". In: *IEEE Transactions on Circuits and Systems for video Technology* 18.11 (2008), pp. 1544–1554.

[6] Benedict Low et al. "A systematic review of collective tactical behaviours in football using positional data". In: *Sports Medicine* 50 (2020), pp. 343–385.

[7] *STATSports*. `https://pro.statsports.com/soccer/`. Accessed: 2024-12-09.

[8] Daniele Conte et al. "Investigating the game-related statistics and tactical profile in NCAA division I men's basketball games". In: *Biology of sport* 35.2 (2018), pp. 137–143.

[9] Franc Erčulj and Erik Štrumbelj. "Basketball shot types and shot success in different levels of competitive basketball". In: *PloS one* 10.6 (2015), e0128885.

[10] Joachim Gudmundsson and Michael Horton. "Spatio-temporal analysis of team sports". In: *ACM Computing Surveys (CSUR)* 50.2 (2017), pp. 1–34.

[11] Bin Li and Xinyang Xu. "Application of artificial intelligence in basketball sport". In: *Journal of Education, Health and Sport* 11.7 (2021), pp. 54–67.

[12] Changjia Tian et al. "Use of machine learning to automate the identification of basketball strategies using whole team player tracking data". In: *Applied Sciences* 10.1 (2019), p. 24.

[13] Mark Harmon et al. "Predicting shot making in basketball learnt from adversarial multiagent trajectories". In: *International Journal of Sport and Health Sciences* 15.11 (2021), pp. 973–983.

[14] Feiyu Xu et al. "Explainable AI: A brief survey on history, research areas, approaches and challenges". In: *CCF International Conference on Natural Language Processing and Chinese Computing*. Springer, Cham. 2019, pp. 563–574.

[15] Zhenni Feng and Yanmin Zhu. "A survey on trajectory data mining: Techniques and applications". In: *IEEE Access* 4 (2016), pp. 2056–2067.

[16] Jae-Gil Lee, Jiawei Han, and Kyu-Young Whang. "Trajectory clustering: a partition-and-group framework". In: *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*. 2007, pp. 593–604.

[17] Aoxue Zang et al. “Towards Predicting Vehicular Data Consumption”. In: *2021 22nd IEEE International Conference on Mobile Data Management (MDM)*. IEEE. 2021, pp. 109–114.

[18] Eleni Chondrodima et al. “Machine learning models for vessel route forecasting: An experimental comparison”. In: *2022 23rd IEEE International Conference on Mobile Data Management (MDM)*. IEEE. 2022, pp. 262–269.

[19] Xinyu Di et al. “Traffic congestion prediction by spatiotemporal propagation patterns”. In: *2019 20th IEEE International Conference on Mobile Data Management (MDM)*. IEEE. 2019, pp. 298–303.

[20] Liang Ge et al. “Temporal graph convolutional networks for traffic speed prediction considering external factors”. In: *2019 20th IEEE International Conference on Mobile Data Management (MDM)*. IEEE. 2019, pp. 234–242.

[21] Tanvir Ahmed et al. “Online risk prediction for indoor moving objects”. In: *2016 17th IEEE International Conference on Mobile Data Management (MDM)*. Vol. 1. IEEE. 2016, pp. 102–111.

[22] Wei Wu et al. “To taxi or not to taxi?-enabling personalised and real-time transportation decisions for mobile users”. In: *2012 IEEE 13th International Conference on Mobile Data Management*. IEEE. 2012, pp. 320–323.

[23] Abdullah Sawas et al. “Tensor methods for group pattern discovery of pedestrian trajectories”. In: *2018 19th IEEE International Conference on Mobile Data Management (MDM)*. IEEE. 2018, pp. 76–85.

[24] Bolei Zhou, Xiaoou Tang, and Xiaogang Wang. “Learning collective crowd behaviors with dynamic pedestrian-agents”. In: *International Journal of Computer Vision* 111 (2015), pp. 50–68.

[25] Shuai Yi, Hongsheng Li, and Xiaogang Wang. “Understanding pedestrian behaviors from stationary crowd groups”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 3488–3496.

[26] Liushuai Shi et al. “SGCN: Sparse graph convolution network for pedestrian trajectory prediction”. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021, pp. 8994–9003.

[27] Zihao Sheng et al. “Graph-based spatial-temporal convolutional network for vehicle trajectory prediction in autonomous driving”. In: *IEEE Transactions on Intelligent Transportation Systems* 23.10 (2022), pp. 17654–17665.

[28] Long Sha et al. “Chalkboarding: A new spatiotemporal query paradigm for sports play retrieval”. In: *Proceedings of the 21st International Conference on Intelligent User Interfaces*. 2016, pp. 336–347.

[29] Toshihiro Tani, Hung-Hsuan Huang, and Kyoji Kawagoe. “Sports play visualization system for American football”. In: *Proceedings of the international multiconference of engineers and computer scientists*. Vol. 1. 2015.

[30] Ulf Brefeld, Jan Lasek, and Sebastian Mair. “Probabilistic movement models and zones of control”. In: *Machine Learning* 108.1 (2019), pp. 127–147.

[31] Qiyun Zhang et al. “Sports match prediction model for training and exercise using attention-based LSTM network”. In: *Digital Communications and Networks* 8.4 (2022), pp. 508–515.

[32] Xinyu Wei et al. “Large-Scale Analysis of Formations in Soccer”. In: *2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE. 2013, pp. 1–8.

[33] Alina Bialkowski et al. "Large-scale analysis of soccer matches using spatiotemporal tracking data". In: *2014 IEEE International Conference on Data Mining*. IEEE. 2014, pp. 725–730.

[34] Qing Wang et al. "Discerning tactical patterns for professional soccer teams: an enhanced topic model with applications". In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 2015, pp. 2197–2206.

[35] Tom Decroos, Jan Van Haaren, and Jesse Davis. "Automatic discovery of tactics in spatio-temporal soccer match data". In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM. 2018, pp. 223–232.

[36] Yasuyuki Kataoka and Peter Gray. "Real-time power performance prediction in tour de France". In: *International workshop on machine learning and data mining for sports analytics*. Springer, Cham. 2019, pp. 121–130.

[37] Kota Itoda, Norifumi Watanabe, and Yoshiyasu Takefuji. "Model-based behavioral causality analysis of handball with delayed transfer entropy". In: *Procedia Computer Science* 71 (2015), pp. 85–91.

[38] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. ""Why should I trust you?" Explaining the predictions of any classifier". In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. ACM. 2016, pp. 1135–1144.

[39] Scott M Lundberg and Su-In Lee. "A unified approach to interpreting model predictions". In: *Advances in neural information processing systems*. Vol. 30. 2017.

[40] Amitojdeep Singh, Sourya Sengupta, and Vasudevan Lakshminarayanan. "Explainable deep learning models in medical image analysis". In: *Journal of Imaging* 6.6 (2020), p. 52.

[41] Richard Meyes et al. "Ablation studies in artificial neural networks". In: *arXiv preprint arXiv:1901.08644* (2019).

[42] Harish G Ramaswamy. "Ablation-CAM: Visual explanations for deep convolutional network via gradient-free localization". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. IEEE. 2020, pp. 983–991.

[43] Bolei Zhou et al. "Learning deep features for discriminative localization". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2921–2929.

[44] Ramprasaath R Selvaraju et al. "Grad-CAM: Visual explanations from deep networks via gradient-based localization". In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 618–626.

[45] Aditya Chattopadhay et al. "Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks". In: *2018 IEEE winter conference on applications of computer vision (WACV)*. IEEE. 2018, pp. 839–847.

[46] Haofan Wang et al. "Score-CAM: Score-weighted visual explanations for convolutional neural networks". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2020, pp. 24–25.

[47] Peiyun Zhong, Di Wang, and Chunyan Miao. "EEG-based emotion recognition using regularized graph neural networks". In: *IEEE Transactions on Affective Computing* (2020).

[48] Charles A Ellis et al. "A Novel Local Ablation Approach For Explaining Multimodal Classifiers". In: *2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE. 2021, pp. 1–6.

[49] MD Zeiler. "Visualizing and Understanding Convolutional Networks". In: *European conference on computer vision/arXiv*. Vol. 1311. 2014.

[50] Ruth C Fong and Andrea Vedaldi. "Interpretable explanations of black boxes by meaningful perturbation". In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 3429–3437.

[51] Ruth Fong, Mandela Patrick, and Andrea Vedaldi. "Understanding deep networks via extremal perturbations and smooth masks". In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2019, pp. 2950–2958.

[52] Siliang Tong, Xueming Luo, and Bo Xu. "Personalized mobile marketing strategies". In: *Journal of the Academy of Marketing Science* 48 (2020), pp. 64–78.

[53] Tianqi Xia et al. "Measuring spatio-temporal accessibility to emergency medical services through big GPS data". In: *Health & place* 56 (2019), pp. 53–62.

[54] Tharindu Fernando et al. "Soft+ hardwired attention: An lstm framework for human trajectory prediction and abnormal event detection". In: *Neural networks* 108 (2018), pp. 466–478.

[55] Nishant Nikhil and Brendan Tran Morris. "Convolutional neural network for trajectory prediction". In: *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. 2018, pp. 0–0.

[56] Toon Bogaerts et al. "A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data". In: *Transportation Research Part C: Emerging Technologies* 112 (2020), pp. 62–77.

[57] Adriana Draghici and Maarten Van Steen. "A survey of techniques for automatically sensing the behavior of a crowd". In: *ACM Computing Surveys (CSUR)* 51.1 (2018), pp. 1–40.

[58] Alina Bialkowski et al. "Identifying team style in soccer using formations learned from spatiotemporal tracking data". In: *2014 IEEE international conference on data mining workshop*. IEEE. 2014, pp. 9–14.

[59] Nazanin Mehrasa et al. "Deep learning of player trajectory representations for team activity analysis". In: *11th mit sloan sports analytics conference*. Vol. 2. 2018, p. 3.

[60] Chien-Yi Chen et al. "Adversarial generation of defensive trajectories in basketball games". In: *2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE. 2018, pp. 1–1.

[61] Stats Perform. *Sport VU*. Accessed: 2024-12-09. 2024. URL: `https://www.statsperform.com/team-performance/basketball/optical-tracking/`.

[62] Wikipedia. *Efficiency (basketball)*. Accessed: 2024-12-09. 2024. URL: `https://en.wikipedia.org/wiki/Efficiency_(basketball)`.

[63] Jaime Sampaio et al. "Exploring game performance in the National Basketball Association using player tracking data". In: *PloS one* 10.7 (2015), e0132894.

[64] Alexander Franks et al. "Characterizing the spatial structure of defensive skill in professional basketball". In: *The Annals of Applied Statistics* 9.1 (2015), pp. 94–121.

[65] Basketball Reference. *Player Efficiency Rating (PER)*. URL: `https://www.basketball-reference.com/about/per.html`.

[66] Football LAB. *Football LAB*. URL: `https://www.football-lab.jp/`.

[67] Jordan L Fox, Aaron T Scanlan, and Robert Stanton. "A review of player monitoring approaches in basketball: current trends and future directions". In: *The Journal of Strength & Conditioning Research* 31.7 (2017), pp. 2021–2029.

[68] Wan-Lun Tsai, Tse-Yu Pan, and Min-Chun Hu. “Feasibility study on virtual reality based basketball tactic training”. In: *IEEE Transactions on Visualization and Computer Graphics* 28.8 (2020), pp. 2970–2982.

[69] Antonia Creswell et al. “Generative adversarial networks: An overview”. In: *IEEE signal processing magazine* 35.1 (2018), pp. 53–65.

[70] *Premier League and Match Insights—Powered by Oracle Cloud: Reimagining the fan experience.* `https://www.oracle.com/premier-league/`. Accessed: 2024-12-09.

[71] George Wehbe et al. “Monitoring neuromuscular fatigue in team-sport athletes using a cycle-ergometer test”. In: *International Journal of Sports Physiology and Performance* 10.3 (2015), pp. 292–297.

[72] Martin Buchheit et al. “Monitoring fitness, fatigue and running performance during a pre-season training camp in elite football players”. In: *Journal of Science and Medicine in Sport* 16.6 (2013), pp. 550–555.

[73] Naotake Inoue and Toyoto Sugiyama. “A Study on Team Building and Game Performance in Soccer”. In: *Journal of Sports Performance Research* 1 (2009), pp. 162–168.

[74] Jungo Oe et al. “Objective Evaluation of Game Performance in Soccer”. In: *Japan Journal of Physical Education, Health and Sport Sciences* 58.2 (2013), pp. 731–736.

[75] Shane Malone et al. “High-speed running and sprinting as an injury risk factor in soccer: Can well-developed physical qualities reduce the risk?” In: *Journal of Science and Medicine in Sport* 21.3 (2018), pp. 257–262.

[76] Tim J Gabbett. “The training—injury prevention paradox: should athletes be training smarter and harder?” In: *British Journal of Sports Medicine* 50.5 (2016), pp. 273–280.

[77] Ricardo Duarte et al. “Intra-and inter-group coordination patterns reveal collective behaviors of football players near the scoring zone”. In: *Human Movement Science* 31.6 (2012), pp. 1639–1651.

[78] Katsuhiro Yokoyama et al. “Social forces for team coordination in ball possession game”. In: *Physical Review E* 97.2 (2018), p. 022410.

[79] Michael Nevett et al. “Changes in Basic Tactics and Motor Skills in an Invasion-Type Game after a 12-Lesson Unit of Instruction”. In: *Journal of Teaching in Physical Education* 20.4 (2012), pp. 352–369.

[80] Stephen Mitchell, Judith Oslin, and Linda Griffin. *Teaching Sport Concepts and Skills: A Tactical Games Approach*. 2nd. Human Kinetics Publishers (UK) Ltd, 2016.

[81] Daniel Memmert and Klaus Roth. “The effects of non-specific and specific concepts on tactical creativity in team ball sports”. In: *Journal of Sport Science* 25 (2007), pp. 1423–1432.

[82] Katsunori Kitano. “Analysis of Neural Circuits Using Transfer Entropy”. In: *Annual Review Neuroscience* (2017), pp. 1–5.

[83] Shohei Hidaka. “Visualization of Information Networks in Multisensory Human Communication”. In: *Journal of the Institute of Electronics, Information and Communication Engineers* 96.12 (2013), pp. 945–950.

[84] Yuki Tenjin and Tomonobu Ozaki. “Analysis of the Interrelation Between SNS and Product Purchasing Using Dynamic Networks with Transfer Entropy”. In: *Proceedings of the Japanese Society for Artificial Intelligence Workshop*. 2015, pp. 13–17.

[85] Hidenobu Sumioka, Yuichiro Yoshikawa, and Minoru Asada. “Discovery of Causal Relationships Using Transfer Entropy and Collaborative Attention Acquisition”. In: *Proceedings of the Robotics and Mechatronics Conference of the Japan Society of Mechanical Engineers*. Vol. 1A2-L08. 2017.

[86] Kazuki Komura, Fujio Toriumi, and Hirotada Ohashi. "Analysis of Investment Indicators Using an Inter-Stock Influence Network Based on Transfer Entropy". In: *Proceedings of the Japanese Society for Artificial Intelligence Workshop on Financial Information*. 2014, pp. 1–6.

[87] Hideaki Ishii, Roberto Tempo, and Er-Wei Bai. "A web aggregation approach for distributed randomized PageRank algorithms". In: *IEEE Transactions on Automatic Control* 57.11 (2012), pp. 2703–2717.

[88] Nadeem Akhtar. "Social network analysis tools". In: *2014 Fourth International Conference on Communication Systems and Network Technologies*. IEEE. 2014, pp. 388–392.

[89] Seiko Sasaki et al. "Comparison of a new wrist-worn accelerometer with a commonly used triaxial accelerometer under free-living conditions". In: *BMC Research Notes* 11.1 (2018), pp. 1–5.

[90] Tatsuya Tamura, Hiroyuki Horino, and Jun Tsuchiya. "Proposal and Examination of Classification Methods for Post-Ball Recovery Attacks in Soccer: Focusing on Fast Breaks and Possession Attacks in the 2012 UEFA European Championship". In: *Journal of Sports Science Research* 12 (2015), pp. 48–55.

[91] Yusuke Sato, Ryo Yamanaka, and Hiroshi Otsuka. "Extraction of Passes from a Soccer Offensive-Defensive Transition Model". In: *Proceedings of the Forum on Information Technology*. Vol. 14. 1. 2015, pp. 145–148.

[92] Jun Sone. "A Kinesiological Study on Instruction for Enhancing Kinesthetic Knowledge of Pass Receivers in Soccer". In: *Journal of Sports and Exercise Studies* 30 (2018), pp. 30–48.

[93] Kazuo Sekine and Katsuya Takanashi. "Cues for Defensive Players in Soccer to Bridge Temporal and Spatial Gaps with Offensive Players". In: *Cognitive Science* 19.2 (2012), pp. 244–248.

[94] Shinya Ito et al. "Extending transfer entropy improves identification of effective connectivity in a spiking cortical network model". In: *PLoS One* 6.11 (2011), e27431.

[95] Toshihiro Nakajima. "Improvement of Physical Education Classes Using Advanced Technology: Efforts to Enhance Teachers' Teaching Skills by Combining Educational Practices and ICT". In: *Journal of Sport Education Studies* 42.1 (2022), pp. 41–43. DOI: 10.7219/jjses.42.1_41.

[96] Divya R Seshadri et al. "Wearable sensors for monitoring the internal and external workload of the athlete". In: *NPJ Digital Medicine* 2.1 (2019), pp. 1–18.

[97] Takeshi Tanaka, Akira Uchiyama, and Hirozumi Yamaguchi. "ML-based Individual Contribution Assessment of Basketball Players from Their Trajectories". In: *2023 24th IEEE International Conference on Mobile Data Management (MDM)*. IEEE. 2023, pp. 55–64.

[98] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. "Convolutional neural networks on graphs with fast localized spectral filtering". In: *Advances in neural information processing systems*. Vol. 29. 2016, pp. 3844–3852.

[99] Justin Gilmer et al. "Neural message passing for quantum chemistry". In: *Proceedings of the 34th International Conference on Machine Learning*. PMLR. 2017, pp. 1263–1272.