Deep reinforcement learning in a racket sport for player evaluation with technical and tactical contexts

NING DING¹, KAZUYA TAKEDA¹ (Senior Member, IEEE), and KEISUKE FUJII¹,²,³ (Member, IEEE)
¹Graduate School of Informatics, Nagoya University, Furo-cho, Chikusa, Nagoya, Aichi, Japan
²RIKEN Center for Advanced Intelligence Project, 744 Motooka, Nishi, Fukuoka, Japan
³JST PRESTO, 7 Gobancho, Chiyoda, Tokyo, Japan

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

This work was supported by JSPS KAKENHI (Grant Numbers 20H04075 and 21H05300) and JST Presto (Grant Number JPMJPR20CA).

ABSTRACT Evaluating the performance of players in a dynamic competition is vital for achieving effective sports coaching. However, a quantitative evaluation of players in racket sports is difficult because it is derived from the integration of complex tactical and technical (i.e., whole-body movement) performances. In this study, we propose a new evaluation method for racket sports based on deep reinforcement learning, which can analyze the motion of a player in more detail, rather than only considering the results (i.e., scores). Our method uses historical data including information related to the tactical and technical performance of players to learn the next-score probability as a Q-function, which is used to value the actions of the players. We leverage long short-term memory model for the learning of Q-function with the poses of the players and the position of the shuttlecock as the input, which are identified by the AlphaPose and TrackNet algorithms, respectively. We verified our approach by comparing various baselines and demonstrated the effectiveness of our method through use cases that analyze the performance of the top badminton players in world-class events.

INDEX TERMS deep reinforcement learning, sports analytics, Q values, action evaluation.

I. INTRODUCTION

With the development of deep learning technologies, computer vision-based deep learning methods have become increasingly important in sports analytics. Among the broad scope of tasks to be addressed, player evaluation is a major task; it has received increasing attention as it can assess and appraise the actions observed in a game to players, coaches, and other staff in order to facilitate decision making (i.e., tactics) and improve technical skills, thus providing a competitive advantage to an individual or a team.

There are two main approaches for player evaluation. The first is to use various statistics to sum up “the total contributions of a player to his/her team” into a number. The second approach is to assign values to the actions performed during a match. In the second approach, traditional methods (e.g., [1]) demonstrate significant limitations as they can only evaluate the actions that directly lead to a score (e.g., shooting), but are unable to evaluate those actions that indirectly lead to score. Recently, Markov models have been used to address this issue, which have the advantage of unified evaluation criteria (actions are evaluated in the same scale by anticipating expected outcomes). These approaches are based on the analysis of event stream data (including optical data) that describe the actions performed in a game. However, in racket sports, the task of action value evaluation of players is almost unexplored because, in such a sport, technical whole-body movements have been evaluated in addition to the tactics.

In racket sports, most previous studies on player evaluation were limited to traditional methods. Shen et al. evaluate a single shot type (lob) in badminton, and defined a good/bad lob on whether the subsequent action of the opponent is a smash and leads to a score [2]. This work was considered in the limited available research; thus, inspired by this idea, we adopt the Markov model approach (i.e., the second approach in the above paragraph) using technical whole-body movement and tactical information for player evaluation in actual games.

With the recent advances in deep learning, deep reinforce-
mement learning (DRL) [3] has been applied in various fields [4]–[7], and has shown significant promise in player evaluation in complex and dynamic environments. A previous method used the Markov model in team sports (e.g., ice hockey) to evaluate the tactical performance, but ignored the effect of technical performance on the value of the action. As racket sports have the characteristic of less participants (single- or double-player games), the result of a game depends largely on individual technical skills. Therefore, when compared to team sports, the specific sports-based technical performance of player(s) should be analyzed through videos. Most previous studies have used active reinforcement learning (RL) to calculate optimal strategies for complex continuous-flow games [8], [9]. Similar to Liu et al. [10], we solve a prediction (not control) problem in the passive learning (fixed policy) setting. We use RL as a behavioral analytic tool for real human agents, not to control artificial agents.

In this study, we propose a player evaluation approach with play contexts in badminton, which leverages historical match data containing tactical and technical information to assign a rating to the actions (e.g., smash) performed by the players in a match. For a given badminton game, we used deep learning methods to extract the features from videos that contain information related to both the tactical and technical performance of the players. Through experiments, we examined the effect of technical and tactical contexts on player evaluation by applying a DRL method on a badminton dataset (videos of the BWF: Badminton World Federation). Our research aims to provide coaches some insight into the influence of the movements of players on the advantages and disadvantages in specific competition situations. Therefore, we analyzed the content of concrete movements and its influence.

In summary, the primary contributions of this study are as follows.

1. We propose a player evaluation method in racket sports based on deep reinforcement learning (DRL) that can analyze the motion of a player in more detail, instead of analyzing the results (i.e., scores).
2. Methodologically, our method leverages historical data including the tactical and technical performance information of players to learn the next-score probability as a Q-function, which is used to value the actions of the players.
3. In the experiment, we verified our approach by comparing various baselines, and confirmed the effectiveness of our method through use cases that analyze the performance of the top badminton players in world-class events.

The remainder of this paper is organized as follows. An overview of the related works is presented in Section II, our method is described in Section III, the experimental results are discussed in Section IV, and this paper is concluded in Section V.

II. RELATED WORK

Evaluation methods in many sports. There are two main approaches employed in evaluating sports players. One involves boiling down the contributions of a player into a single number. Some well-known examples are Wins Above Replacement (WAR) [11] in baseball and Player efficiency rating (PER) [12] in basketball. The other approach involves quantifying the value of the action of a player using the Markov model. Recently, the latter approach has been applied in various sports. Cervone et al. proposed Expected Possession Value (EPV) to evaluate the decisions and actions of players using spatiotemporal tracking data in basketball [13]. Liu et al. introduced Game Impact Metric (GIM) to aggregate the values of the actions of players in ice hockey [10]. In soccer, a method of Valuing Actions by Estimating Probabilities (VAEP) [14] and its variants [15] has been proposed. A survey on team sports was recently conducted in [16].

Based on reinforcement learning frameworks, certain papers were published on team sports using inverse planning methods, which estimate the action model or reward function from the observed data using statistical learning techniques. For example, the state-action value function (Q-function) of a player was estimated using a recurrent neural network [10], [17], which was interpreted using a linear model tree [18]. To evaluate the shooting action of players, researchers investigated the expected possession value [19]–[21], and the value of the space [22], [23] by extending a Voronoi diagram [24]. In team games, DRL was used for estimating the quality of defensive actions in ball-screen defense in basketball [25]. Compared to these studies on invasive sports, we propose a DRL framework for racket sports utilizing technical (i.e., pose) information and tactical contexts.

Quantitative analysis in racket sports. In racket sports, there is a broad range of tasks to address. Most of the related research focused on the detection/tracking of players and sports equipment (e.g. ball) [26]–[28], stroke classification [29]–[31], and match outcome prediction [32], [33]. Despite the popularity of racket sports, only a limited number of works offer computer vision-based solutions for the task of player evaluation. Certain previous works [34] manually designed evaluation formula for players (i.e., first approach in the above paragraph). Pfeiffer et al. were the first to adopt the Markov model approach for performance diagnosis in table tennis [35]. McGarry and Franks used the Markov model for explaining the championship performance in squash [36]. However, these studies discretized the coordinates of location and time, which resulted in the loss of information and failed to generalize the unobservable parts of the state space.

For stroke evaluation, traditional methods require sports analysts to examine the videos repeatedly, and evaluate the stroke performance based on their knowledge. Deep learning-based computer vision approaches extract stroke features from videos [35], [37], [38] to characterize and model the competition process. Wang et al. [39] integrated the knowledge of analysts, and trained a classifier that learned the evaluation of analysts to obtain quantified evaluation results; however, their approach using domain knowledge (table tennis) varies from our general approach for modeling racket
sports. Recently, Wang et al. [40] introduced a badminton language to describe the process, and predict the win probability in a rally. Compared to reinforcement learning methods, their method cannot estimate the value of each stroke directly.

III. PROPOSED METHOD
In this section, we describe the proposed method for estimating the action value of a player. The overview of our approach is illustrated in Fig. 1.

Badminton is a competitive sport, where the winner of a match is determined based on the best performance out of three games and each game is played for 21 points. A rally starts with a serve and ends when the point is won. To describe a badminton match, we considered a rally as the analysis unit. The course of the rally can be described as the transition from one state to another. A rally in badminton comprises a sequence of strokes with outcomes.

By using the video data of badminton games, we first segmented each game in a match into several rallies. For each rally, we estimated the XY-coordinate values (17 body parts in the COCO (Common Objects in Context) dataset) for joint positions of nose, eyes, ears, shoulders, elbows, wrists, hips, knees, and ankles using AlphaPose [41]—a popular high precision multi-person body-pose estimation system, and detected the shuttlecock position using TrackNet [42]—an object tracking network that has been proven to exhibit a decent tracking capability in games with high-speed balls such as badminton.

As a preprocessing procedure, joint positions that were not properly estimated owing to an overlap were annotated through the COCO annotator. Moreover, we assumed that the midpoint of the two ankles indicates the position of the player, and the pose represents the coordinate values relative to the position of the player. For the pose of a left-handed player, we reversed the corresponding relative coordinate values.

We combined the outputs as an input feature vector of a DRL model, applied the DRL method to estimate the Q-values, and finally, obtained the action value from the Q-values for evaluating the performance of a player. According to the position of the players relative to the camera, we divided the two players in a game into the front and back players. We considered the player closer to camera as the front player; conversely, the player farther from the camera is the back player (we sometimes denote them as front and back, respectively).

A. FORMULATION
We utilize a reinforcement learning framework, specifically based on the recent sports-related work [10]. The reward $R$ specifies the player who wins a point at the end of a rally. In this study, we generalized the strokes to nine types, namely serve, drop, smash, clear, lift, drive, block, net kill, and net shot. Fig. 2 illustrates the different stroke types. Action $a_t$ is one of the stroke types.

![Figure 1: Overview of our approach. Pose estimation of players and shuttlecock position detection using AlphaPose [41] and TrackNet [42], followed by a DRL model for estimating the Q-function. Given a Q-function, the action value is defined as the change in Q-value due to the action.](image1)

![Figure 2: Illustration of different stroke types. The trajectories of (1) serve, (2) drop, (3) smash, (4) clear, (5) lift, (6) drive, (7) block, (8) net kill, (9) net shot are shown.](image2)

To describe the state, we considered the feature vector at each hit time (the moment when the racket contacts the shuttlecock) as state representation $s_t$ at time step $t$.

The Q-function $Q(s, a)$, represents the conditional probability that the front resp. back player winning the point at the end of the rally such that

$$Q_{front/back}(s, a) = P(point = 1|s_t = s, a_t = a).$$

The Q-function computes the expected rewards for an action taken in a given state. Different Q-functions can be used to study different outcomes of interest, such as goals and penalties [43]. In a variation from a previous study [35] that used “point” and “fault” as the expected outcomes, we used the next-score probability as the Q-function. The advantages are as follows. 1) Compared to the outcome of a rally (“point” or “fault”), the next-score probability function is highly interpretable, as it models the probability of an event. 2) It can provide coaches with a more detailed overview of player performance during a rally.

B. LEARNING Q FUNCTION
Fig. 3 shows the architecture of our DRL model. We constructed dynamic two-layer long short-term memory (LSTM) networks to learn the Q-function and estimate the Q-values.
The networks take a sequence of states $s_t$ and actions $a_t$ at the moment the player hits the shuttlecock (hit frame) as the input. We used this model to simultaneously evaluate both the front and back players in a given rally.

The LSTM networks have an input layer with 256 nodes, a hidden layer with 256 nodes, and a dense output layer with 3 output nodes. The three output nodes, $Q_{\text{Back}}(s, a)$, $Q_{\text{Front}}(s, a)$, and $Q_{\text{Rally\_end}}(s, a)$, represent the probability that the front/back player wins the next point according to the present state and action and the probability that a rally ends according to the present state and action. The LSTM networks require four types of input data for model training, namely the XY-coordinates of the positions of the front and back players, both their poses, and the position of the shuttlecock, and the action at each hit time.

We applied state–action–reward–state–action, which is an on-policy reinforcement learning algorithm for estimating Q-values. The Q-value for a state-action is updated using the following equation:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[ R + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right],$$

(3)

where $s_{t+1}$ and $a_{t+1}$ denote the state and action at time step $t+1$. $\alpha$ is the learning rate, and $\gamma$ is the discount factor. Instead of tabular learning, we considered the neural networks as function approximators.

Our network is trained with loss $L$ as:

$$L = \frac{1}{n} \sum_{t=1}^{n} \left( Q(s_t, a_t) - R - \gamma Q(s_{t+1}, a_{t+1}) \right)^2.$$

(4)

The network is trained using the Adam optimizer with a learning rate of $10^{-4}$ for 90 epochs. The hyperparameters were set to $\gamma = 0.3$.

**C. STROKE EVALUATION**

For each action of the front/back player in a rally, action value $A$ can be computed as:

$$A_{t+1} = Q_{\text{player}}(s_{t+1}, a_{t+1}) - Q_{\text{player}}(s_t, a_t).$$

(5)

We set the probability that the front/back player will win the next point (Q-value) before service as 0. We calculated...
the action value of each stroke type and the corresponding number of each stroke type that a player takes in each game, and calculated the average action value of each stroke type performed by a player in a game, as shown in Fig. 4.

We present a typical example to demonstrate the effectiveness of our approach. The graph on the left in Fig. 4 shows the dynamic changes in the Q-values of a rally in a match between Lin (back player) and Lee (front player), which was conducted on March 17, 2018. The back player Lin won this rally in the end. The figure plots the values of the three output nodes. The graph on the right in Fig. 4 shows the action value of each stroke performed by the front and back players. According to the bar chart, the fourth stroke, which was a “smash” performed by Lin, has the greatest action value, and the fifth stroke, which was a “block” performed by Lee, has a minor action value. The result of the action value is consistent with our intuitive feelings from the video that the “smash” performed by Lin was significantly powerful, and Lee failed to intercept the shuttle through the “block” stroke.

Moreover, this figure provides more insights on the rally. The intuitive feelings gained from the video may make the coach erroneously attribute the loss of a point to the poor defensive “block” performed by Lee. However, our result shows the action value of the “block” is positive, albeit a small one. Conversely, the third stroke, which was a “clear” stroke performed by Lee, has a negative action value; after this stroke, Lin seized the chance to attack using the offensive “smash” action. We examined the video and identified that Lee returned the shuttle to a position closer to Lin with the “clear” action; consequently, Lin has sufficient time to jump and attack with a “smash.” The figure reflects that the previous actions in a rally have an indirect impact on the final result. Therefore, we can conclude that the result of our approach can provide useful clues for further analysis over the course of multiple strokes in a rally.

IV. EXPERIMENTS
A. DATASETS
We collected videos from the BWF TV channel. We ignored the cross-view problem, and only selected videos captured from a single view (as shown in Fig. 1). Our constructed dataset provides information regarding game contexts and player actions for the 2018-2020 BWF season, which contains 21 matches, covering 22 players and 320 rallies.

B. VERIFICATION OF OUR METHOD
We evaluated the design choices of our approach in terms of input components by comparing the performance of the LSTM networks with different inputs.

As presented in Table 1, a model trained with all the input components (shuttle position, player position, and player poses) achieves the best performance, indicating that all the input components contribute in estimating an accurate Q-function.

Next, we examined the effect of the stroke type on our model. As the action is a primary element in reinforce-

| TABLE 1: Comparison of each input feature by eliminating the different input features from the total input. We report loss function $\mathcal{L}$ according to Eq. (4), which was evaluated on a badminton dataset. |
|---|---|---|---|---|
| Error (Shuttle position) | Error (Player position) | Error (Player pose) | Full |
| $\mathcal{L}$ | 0.0289 | 0.0456 | 0.0506 | 0.0261 |

C. CHARACTERISTICS OF THE PLAYERS IDENTIFIED USING OUR METHOD
The matches considered here include one final, two semifinals, and one quarter-final men’s singles matches in the 2018 BWF World Tour, which are the matches in the All England Open Tournament between Lin Dan and Shi Yuqi, Huang Yuxiang and Lin Dan, Son Wan-ho and Shi Yuqi, and Lin Dan and Lee Chong Wei. We examined the performance of the players from the perspective of the average action value of each stroke type performed by a player in a match (Fig. 5).

For example, we identified that, when Lin was playing against Huang and Lee (Figs. 5 (B) and (D)), the strokes performed by Lin were superior to those by his opponents. Lin had fewer stroke types whose action value were substantially below zero, especially when he was playing against Lee. His offensive strokes such as “smash” and “net shot” were better than those of Lee. However, when playing against Shi (Fig. 5 (A)), Lin did not display significant advantages, which would explain why he lost the final match. During the match between Shi and Son (Fig. 5 (C)), Shi showed superior performance in offensive strokes such as “smash” and “net shot.” The summary of all player action values in the 2018–2020 BWF Tour was indicated in Table 2.

D. RELATIONSHIP WITH SCORE
Owing to the cross-view scenes in badminton videos, we could only use a portion of the rallies in each match. Therefore, to further demonstrate the effectiveness of our approach, we examined the relationship between average (maximum) action values and the score, as shown in Figs. 6 (A) and (B). Here, the average (maximum) action values refer to the average value of the maximum value of the actions performed by a player in each rally over several rallies, and the score indicates the number of points scored by the front/back player in a match.

Spearman’s rank correlation coefficient $\rho$ was applied to quantify the correlation, because we also examined the relationship with the rank of the players, as described in the next subsection. Fig. 6 (A) shows that $\rho = 0.150$ ($p > 0.05$),
FIGURE 4: Typical example of the analysis in a rally. Left: For each rally, the graph shows the probability that the front/back player scores the next point, as estimated by the DRL model. A greater action value leads to major changes in the scoring chance. The network can also capture smaller changes associated with every action under different technical and tactical contexts. Right: Action value of each stroke performed by the front/back player.

TABLE 2: Player action values in the 2018–2020 BWF Tour

| Game | Players | Average action value | Average (maximum) action value | Point | Rank |
|------|---------|----------------------|-------------------------------|-------|------|
| 1    | LIN D. vs LEE C. W. | 0.0392 vs 0.0387 | 0.2218 vs 0.2317 | 10 vs 8 | BWF2018 11 vs 49 |
| 2    | NG K. L. vs NISHIMOTO | 0.0169 vs 0.1074 | 0.1634 vs 0.3628 | 5 vs 9 | BWF2018 16 vs 9 |
| 3    | LIN D. vs SHI Y. Q. | 0.0519 vs 0.0630 | 0.1829 vs 0.1871 | 5 vs 6 | BWF2018 11 vs 4 |
| 4    | CHOU T. C. vs TSUNEYAMA | 0.0814 vs -0.0030 | 0.2900 vs 0.2099 | 4 vs 3 | BWF2020 6 vs 25 |
| 5    | LEE Z. J. vs CHRISTIE | 0.0414 vs 0.0432 | 0.2147 vs 0.0974 | 16 vs 3 | BWF2020 10 vs 19 |
| 6    | CHOU T. C. vs AXELSEN | 0.0802 vs 0.0171 | 0.1530 vs 0.2073 | 3 vs 3 | BWF2018 1 vs 20 |
| 7    | HUANG Y. X. vs LIN D. | 0.0512 vs 0.0482 | 0.2122 vs 0.2678 | 9 vs 15 | BWF2018 17 vs 11 |
| 8    | LEE D. K. vs LEE C. W. | 0.0785 vs 0.0354 | 0.2374 vs 0.2603 | 6 vs 8 | BWF2018 30 vs 49 |
| 9    | VITTINGHUS vs LIN D. | 0.0315 vs 0.0997 | 0.2059 vs 0.1905 | 5 vs 10 | BWF2018 23 vs 11 |
| 10   | LEVERDEZ vs KIDAMBI | 0.0379 vs 0.0940 | 0.1555 vs 0.2333 | 12 vs 9 | BWF2018 24 vs 14 |
| 11   | LEE Z.J. vs AXELSEN | 0.0594 vs 0.2291 | 0.2676 vs 0.2883 | 4 vs 5 | BWF2020 10 vs 1 |
| 12   | CHOU T.C. vs ANTONSEN | 0.0468 vs 0.0571 | 0.2251 vs 0.1094 | 2 vs 2 | BWF2020 6 vs 2 |
| 13   | LEE Z.J. vs CHEN L. | 0.0829 vs 0.0711 | 0.2493 vs 0.0884 | 2 vs 2 | BWF2020 10 vs 28 |
| 14   | LU G.Z. vs AXELSEN | 0.1588 vs 0.1097 | 0.3024 vs 0.2649 | 10 vs 7 | BWF2019 19 vs 5 |
| 15   | SON W.H. vs SHI Y.Q. | 0.0431 vs 0.0237 | 0.1395 vs 0.1667 | 2 vs 6 | BWF2018 46 vs 23 |
| 16   | ANTONSEN vs SHI Y.Q. | 0.0126 vs 0.1541 | 0.1274 vs 0.2332 | 5 vs 10 | BWF2019 27 vs 4 |
| 17   | SHI Y.Q. vs AXELSEN | 0.0184 vs -0.0356 | 0.1262 vs 0.2114 | 5 vs 11 | BWF2020 7 vs 1 |
| 18   | CHEN L. vs GINTING | 0.0627 vs 0.1483 | 0.1317 vs 0.2719 | 7 vs 15 | BWF2019 4 vs 6 |
| 19   | MOMOTA vs GINTING | 0.0414 vs 0.1211 | 0.1646 vs 0.2032 | 15 vs 11 | BWF2019 1 vs 6 |
| 20   | GINTING vs ANTONSEN | 0.0216 vs 0.0292 | 0.1141 vs 0.1566 | 6 vs 4 | BWF2019 6 vs 7 |
| 21   | MOMOTA vs ANTONSEN | 0.0608 vs 0.0489 | 0.2121 vs 0.1994 | 20 vs 20 | BWF2019 1 vs 7 |

indicating that there is no correlation between the score and average action value. Fig. 6 (B) shows that $\rho = 0.312 (p < 0.05)$, which reflects a weak positive monotonic correlation between the score and average (maximum) action value.

The results suggest that the average (maximum) action value could be associated with the score, and if the maximum action value of a player in a rally is greater than that of his opponent, the action can lead to obtaining the score of the rally.

**E. RELATIONSHIP WITH THE RANK OF THE PLAYER**

We also examined the relationship between the rank of the player and average (maximum) action values, as shown in Figs. 6 (C) and (D). Here, the average action values refer to the average value of all actions performed by a player over several rallies. For the rank of a player, we used the official ranking data of the BWF for the front/back player. Fig. 6 (C) shows that $\rho = -0.353 (p < 0.05)$, which reflects a weak negative monotonic correlation between the rank of the player and average action value. Fig. 6 (D) shows that $\rho = -0.048 (p > 0.05)$, which indicates that there is no correlation between the rank of the player and average (maximum) action value.

The results suggest that average action value can be associated with the rank of the player, and higher ranked players tend to perform actions with higher action values.

**F. CORRELATION COMPARISON WITH BASELINE**

No previous studies have used the DRL model to evaluate
FIGURE 5: Average action value of each stroke types in four badminton finals: (A) Final, (B) Semi-Final, (C) Semi-Final, (D) Quarter-Final.

each stroke in racket sports and compared with the similar method in team sports of ice hockey [10], our study took into account the importance of technical performance in single-racket sports. To evaluate our method, we eliminated the information of technical performance (player pose) and used it as a baseline.

Here, we computed the same Spearman’s rank correlation coefficient $\rho$ as above regarding the baseline model. We summarize the results in Table 3, which shows that there was no correlation between player evaluation metrics and standard success for baseline model ($ps > 0.05$). Compared to the baseline, our method (full model) shows a weak monotonic correlation in the two success measures. The correlation results suggest some important information in our full model (pose information) to evaluate the performance of the players.

V. CONCLUSIONS

In this study, we applied the DRL method to the task of player evaluation in racket sports (badminton). We developed a framework to estimate the action value of a player from an input video. We verified our approach by comparing various baselines, and validated the effectiveness of our method through use cases that analyzed the performance of top badminton players in world-class events. We also discovered valuable insights into the correlation between the action value and the score/rank. A badminton dataset with ground truth annotations of both stroke type and game scores is also provided if the paper will be ready for the publication.

In our future work, we plan to improve the framework (e.g., modeling of players separately for player-specific analysis), and extend our framework to other racket sports such as tennis. We also believe that more accurate 3D feature extraction can help in modeling different sports.

ACKNOWLEDGMENT

We would like to thank Tatsuya Yoshikawa for his valuable comments on this work.

REFERENCES

[1] Jake Ensum, Richard Pollard, and Samuel Taylor. Applications of logistic regression to shots at goal at association football. In Science and football: the proceedings of the Fifth World Congress on Science and Football, page 214, 2005.
[2] Leijun Shen, Hui Zhang, Min Zhu, Jia Zheng, and Yawei Ren. Measurement and performance evaluation of lob technique using aerodynamic model in badminton matches. In International Symposium on Computer Science in Sport, pages 53–58. Springer, 2019.
[3] Richard S Sutton and Andrew G Barto. Reinforcement learning: An introduction. MIT press, 2018.
[4] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. nature, 529(7587):484–489, 2016.
[5] Ahmad EL Sallab, Mohammed Abdou, Etienne Perot, and Senthil Yogaman. Deep reinforcement learning framework for autonomous driving. Electronic Imaging, 2017(19):70–76, 2017.
[6] Ying Liu, Brent Logan, Ning Liu, Zhiyuan Xu, Jian Tang, and Yangzhi Wang. Deep reinforcement learning for dynamic treatment regimes on medical registry data. In 2017 IEEE International Conference on Healthcare Informatics (ICHI), pages 380–385. IEEE, 2017.
[7] Shixiang Gu, Ethan Holly, Timothy Lillicrap, and Sergey Levine. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. In 2017 IEEE international conference on robotics and automation (ICRA), pages 3389–3396. IEEE, 2017.
[8] Matthew Hausknecht and Peter Stone. Deep recurrent q-learning for partially observable mdps. In 2015 aaai fall symposium series, 2015.
[9] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. nature, 518(7540):529–533, 2015.
[10] Guiiiang Liu and Oliver Schulte. Deep reinforcement learning in ice hockey for context-aware player evaluation. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18, pages 3442–3448. International Joint Conferences on Artificial Intelligence Organization, 7 2018.
[11] Tobias Gerstenberg, Tomer Ullman, Max Kleinman-Weiner, David Lagudo, and Josh Tenenbaum. Wins above replacement: Responsibility attributions as counterfactual replacements. In Proceedings of the Annual Meeting of the Cognitive Science Society, volume 36, 2014.
FIGURE 6: **Upper** (A, B) Correlation between the score and average action value/average (maximum) action value. **Lower** (C, D) Correlation between the rank of the player and average action value/average (maximum) action value.

TABLE 3: Spearman’s correlation coefficient values between standard success measures and our methods including a baseline. Bold value indicates a statistically significant correlation ($p < 0.05$).

| Model                  | Score Rank | Full model Score Rank |
|------------------------|------------|-----------------------|
| Player pose (baseline) |            |                       |
| Average action value   | $-0.154$   | $-0.122$              |
| Average (maximum) action value | $-0.061$   | $0.163$               |
| Full model             |            |                       |
| Score                  | $0.150$    | $0.342$               |
| Rank                   | $0.312$    | $0.764$               |

[12] Douglas L Cerynik, Gabriel E Lewullah, Brian C Joves, Michael P Palmer, and James A Tom. Outcomes of microfracture in professional basketball players. Knee Surgery, Sports Traumatology, Arthroscopy, 17(9):1135–1139, 2009.

[13] Dan Cervone, Alexander D’Amour, Luke Bornn, and Kirk Goldsberry. Pointwise: Predicting points and valuing decisions in real time with nba optical tracking data. In Proceedings of the 8th MIT Sloan Sports Analytics Conference, Boston, MA, USA, volume 28, page 3, 2014.

[14] Tom Decroos, Lotte Bransen, Jan Van Haaren, and Jesse Davis. Actions speak louder than goals: Valuing player actions in soccer. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pages 1851–1861, 2019.

[15] Kosuke Toda, Masakiyo Teranishi, Keisuke Kushiro, and Keisuke Fujii. Evaluation of soccer team defense based on prediction models of ball recovery and being attacked. arXiv preprint arXiv:2103.09627, 2021.

[16] Keisuke Fujii. Data-driven analysis for understanding team sports behaviors. Journal of Robotics and Mechatronics, 33(3):505–514, 2021.

[17] Guiiliang Liu, Yudong Luo, Oliver Schulte, and Tarak Kharrat. Deep soccer analytics: learning an action-value function for evaluating soccer players. Data Mining and Knowledge Discovery, 34(5):1531–1559, 2020.

[18] Xiangyu Sun, Jack Davis, Oliver Schulte, and Guiiliang Liu. Cracking the black box: Distilling deep sports analytics. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020.

[19] Dan Cervone, Alexander D’Amour, Luke Bornn, and Kirk Goldsberry. Pointwise: Predicting points and valuing decisions in real time with nba optical tracking data. In Proceedings of the MIT Sloan Sports Analytics Conference, 2014.

[20] Daniel Cervone, Alex D’Amour, Luke Bornn, and Kirk Goldsberry. A multiresolution stochastic process model for predicting basketball possession outcomes. Journal of the American Statistical Association, 111(514):585–599, 2016.

[21] Javier Fernandez, Luke Bornn, and Dan Cervone. Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. In Proceedings of the MIT Sloan Sports Analytics Conference, 2019.

[22] Dan Cervone, Luke Bornn, and Kirk Goldsberry. NBA court realty. In Proceedings of the MIT Sloan Sports Analytics Conference, 2016.

[23] Javier Fernandez and Luke Bornn. Wide open spaces: A statistical technique for measuring space creation in professional soccer. In Proceedings of the MIT Sloan Sports Analytics Conference, 2018.

[24] Tsyoshi Taki and Jun-Ichi Hasegawa. Visualization of dominant region
in team games and its application to teamwork analysis. In Proceedings of Computer Graphics International, pages 227–235, 2000.

[25] Jixuan Wang, Yan Fax, Jonathan Skaza, Nick Linke, Satinder Singh, and Jenna Wiens. The advantage of double: a deep reinforcement learning approach to studying the double team in the nba. arXiv preprint arXiv:1803.02940, 2018.

[26] Roman Voeikov, Nikolay Falaleev, and Ruslan Baikulov. Timet: Real-time temporal and spatial video analysis of table tennis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pages 884–885, 2020.

[27] Nur Arzmina Rahmad, Nur Anis Jasmin Sufri, Nurul Hamizah Muzamil, and Muhammad Amir As’ari. Badminton player detection using faster region convolutional neural network. Indonesian Journal of Electrical Engineering and Computer Science, 14(3):1330–1335, 2019.

[28] Fei Yan, W Christmas, and Josef Kittler. A tennis ball tracking algorithm for automatic annotation of tennis match. In British machine vision conference, volume 2, pages 619–628. University of Surrey, 2005.

[29] Hitesh Shah, Prakash Chokalingam, Balamohan Paluri, Nalin Pradeep, and Balasubramanian Raman. Automated stroke classification in tennis. In International conference image analysis and recognition, pages 1128–1137. Springer, 2007.

[30] Kaustubh Milind Kulkarni and Sucheth Shenoy. Table tennis stroke recognition using two-dimensional human pose estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 4576–4584, 2021.

[31] Wei-Ta Chu and Samuel Situmeang. Badminton video analysis based on spatiotemporal and stroke features. In Proceedings of the 2017 ACM on international conference on multimedia retrieval, pages 448–451, 2017.

[32] T Barnett, A Brown, and S Clarke. Developing a model that reflects outcomes of tennis matches. In Proceedings of the 8th Australasian Conference on Mathematics and Computers in Sport, Coolangatta, Queensland, pages 3–5, 2006.

[33] Manoj Sharma, Naresh Kumar, Pardeep Kumar, et al. Badminton match outcome prediction model using naive bayes and feature weighting techniques. Journal of Ambient Intelligence and Humanized Computing, 12(8):8441–8455, 2021.

[34] LIEVEN Vergauwen, ARTHUR J Spaepen, Johan Lefevere, and Peter Hespel. Evaluation of stroke performance in tennis. Medicine and science in sports and exercise, 30(8):1281–1288, 1998.

[35] Mark Pfeiffer, Hui Zhang, and Andreas Hohmann. A markov chain model of elite table tennis competition. International Journal of Sports Science & Coaching, 5(2):205–222, 2010.

[36] Tim McGarry and Ian M Franks. Development, application, and limitation of a stochastic markov model in explaining championship squash performance. Research quarterly for exercise and sport, 67(4):406–415, 1996.

[37] Jiachen Wang, Kejian Zhao, Dazhen Deng, Anqi Cao, Xiao Xie, Zheng Zhou, Hui Zhang, and Yingcai Wu. Tac-simur: Tactic-based simulative visual analytics of table tennis. IEEE transactions on visualization and computer graphics, 26(1):407–417, 2019.

[38] Sebastian Wenninger and Martin Lames. Markov simulation by numerical derivation in table tennis. In Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS), pages 161–169. Springer, 2016.

[39] Jiachen Wang, Dazhen Deng, Xiao Xie, Xinhuai Shu, Yuxuan Huang, Lewen Cai, Hui Zhang, Minling Zhang, Zhihua Zhou, and Yingcai Wu. Tac-validator: Knowledge-based stroke evaluation in table tennis. In The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2021.

[40] Wei-Yao Wang, Teng-Fong Chan, Hui-Kuo Yang, Chih-Chuan Wang, Yao-Chung Fan, and Wen-Chih Peng. Exploring the long short-term dependencies to infer shot influence in badminton matches. arXiv preprint arXiv:2109.06431, 2021.

[41] Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, and Cewu Lu. RMPE: Regional multi-person pose estimation. In ICCV, 2017.

[42] Yu-Chuan Huang, I-No Liao, Ching-Hsuan Chen, Ts-I Ul Ik, and Wen-Chih Peng. Tracknet: a deep learning network for tracking high-speed and tiny objects in sports applications. In 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pages 1–8. IEEE, 2019.

[43] Kurt Routley and Oliver Schulte. A markov game model for valuing player actions in ice hockey. PhD thesis, 2015.

NING DING received her B.E. degree in electronic information science and technology from Guangdong University of Technology, Guangzhou, China, in 2017 and her M.E. degree in information architecture from Waseda University, Kitakyushu, Japan, in 2021. She is currently pursuing her Ph.D. degree in intelligent systems at Nagoya University, Nagoya, Japan. Her research interests include interpretable sports action evaluation and AI coach system.

KAZUYA TAKEDA received his B.E., M.E., and Ph.D. degrees from Nagoya University. He worked at Advanced Telecommunication Research Laboratories and at KDD R&D Laboratories. In 1995, he joined Nagoya University, where he started a research group for signal processing applications. He is currently a Professor at the Graduate School of Informatics and the Green Mobility Collaborative Research Center, Nagoya University. His research interest involves behavior signal processing, including driving behavior.

KEISUKE FUJII received his B.S., M.S., and Ph.D. degrees from Kyoto University in 2009, 2011 and 2014, respectively. After his work as a postdoctoral fellow and research scientist at Nagoya University and RIKEN Center for Advanced Intelligence Project in Japan, he joined Nagoya University. He is currently an associate professor at the Graduate School of Informatics. His research interests are interdisciplinary studies among machine learning, behavioral sciences, and sports sciences.