

Received March 19, 2022, accepted May 5, 2022, date of publication May 16, 2022, date of current version May 26, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3175314

# Deep Reinforcement Learning in a Racket Sport for Player Evaluation With Technical and Tactical Contexts

NING DING<sup>1</sup>, KAZUYA TAKEDA<sup>1</sup>, (Senior Member, IEEE),  
AND KEISUKE FUJII<sup>1,2,3</sup>, (Member, IEEE)

<sup>1</sup>Graduate School of Informatics, Nagoya University, Nagoya, Aichi 464-8601, Japan

<sup>2</sup>RIKEN Center for Advanced Intelligence Project, Fukuoka 819-0395, Japan

<sup>3</sup>JST PRESTO, Tokyo 102-0076, Japan

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

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 20H04075 and Grant 21H05300, and in part by the Japan Science and Technology Agency (JST) Presto under Grant 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 a 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])

The associate editor coordinating the review of this manuscript and approving it for publication was Guangcun Shan<sup>1</sup>.

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 to be 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 stroke 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. The Q-function in the Markov

model has the advantage of evaluating all actions (all stroke types) on the same scale by looking ahead to the outcomes. 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 reinforcement 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 their 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

### A. 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 [13], [19], [20], and the value of the space [21], [22] by extending a Voronoi diagram [23]. In team games, DRL was used for estimating the quality of defensive actions in ball-screen defense in basketball [24]. Compared to these studies on invasive sports, we propose a DRL framework for racket sports utilizing technical (i.e., pose) information and tactical contexts.

### B. QUANTITATIVE ANALYSIS IN RACKET SPORTS

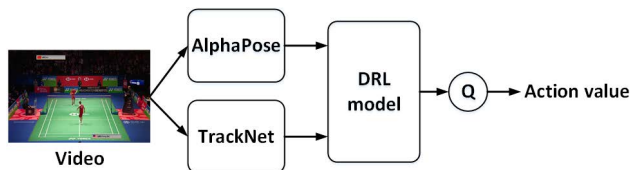
In racket sports, there is a broad range of tasks to address. Most of the related research focused on the detection/tracking of players and sports equipment (e.g. ball) [25]–[27], stroke classification [28]–[30], shuttlecock position prediction [31], and match outcome prediction [32], [33]. Despite the 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.



**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.

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.

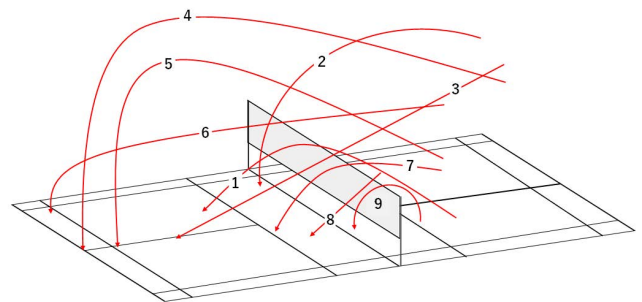
As a preprocessing procedure, joint positions that were not properly estimated owing to an overlap were annotated through the COCO annotator. Moreover, we assumed that the

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 the 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 utilized 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.



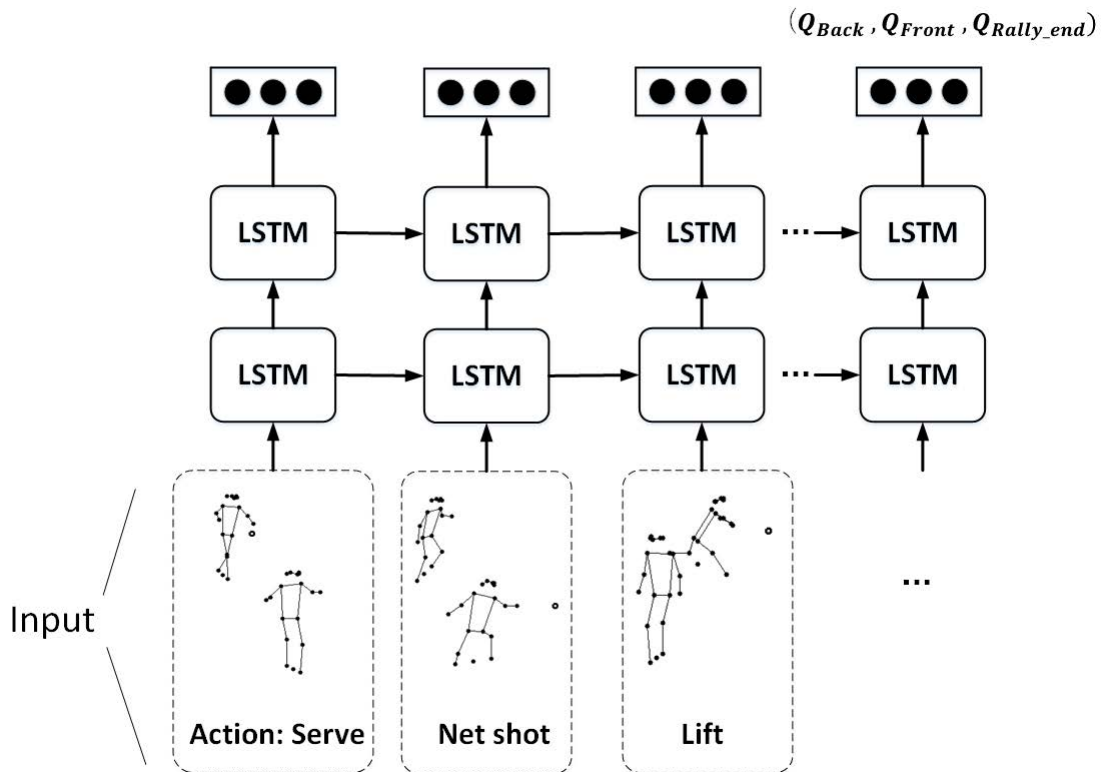
**FIGURE 2.** Illustration of different stroke types. The trajectories of (1) serve, (2) drop, (3) smash, (4) clear, (5) lift, (6) drive, (7) block, (8) net kill, (9) net shot are shown.

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 wins a point at the end of the rally such that

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

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



**FIGURE 3.** Architecture of the DRL model: two layers of LSTM networks. The input is a combination of feature vectors of XY-coordinates of the positions of the front and back players, their poses, the position of the shuttlecock, and the action at each hit time.

event. 2) It can provide coaches with a more detailed overview of player performance during a rally.

### B. LEARNING THE 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,  $Q_{Back}(s, a)$ ,  $Q_{Front}(s, a)$ , and  $Q_{Rally\_end}(s, a)$ , 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, the position of the shuttlecock, and the action.

In this study, the Q-function is learned via a neural network, which is called a function approximation approach. We approximate Q-function by the LSTM networks, we denote LSTM networks as  $q$ , then the approximation can

be written as:

$$Q(s, a) \approx q(s, a; w), \quad (2)$$

where LSTM networks are parameterized by  $w$ . We use the positions of the front and back players, both their poses, and the position of the shuttlecock as the state  $s$ , and the action  $a$  in  $q(s, a; w)$ .

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:

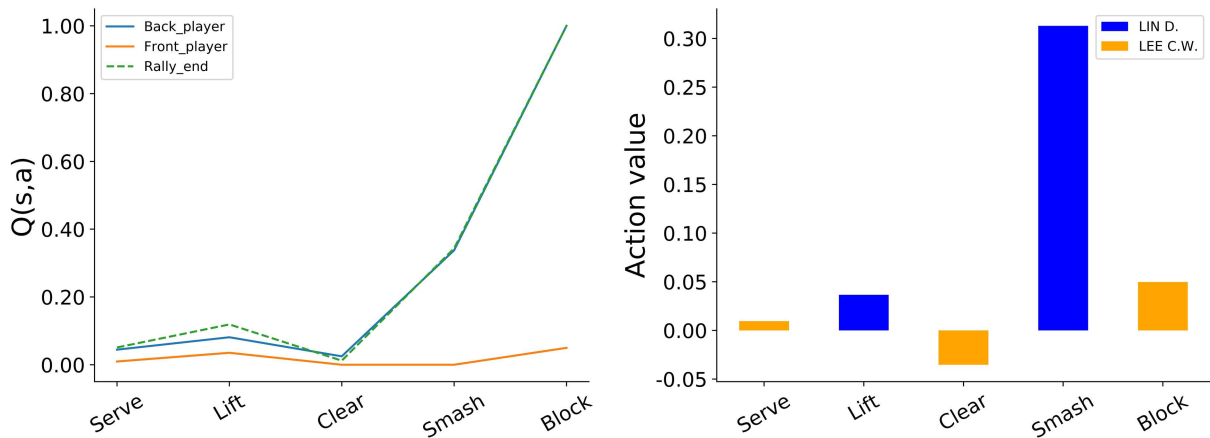
$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [R + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)], \quad (3)$$

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  $\mathcal{L}$  as:

$$\mathcal{L} = \frac{1}{n} \sum_{t=1}^n (Q(s_t, a_t) - R - \gamma Q(s_{t+1}, a_{t+1}))^2. \quad (4)$$

The network is trained using the Adam optimizer with a learning rate of  $10^{-5}$  for 90 epochs. The hyperparameters were set to  $\gamma = 0.3$ .



**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.

### C. STROKE EVALUATION

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

$$A_{t+1} = Q_{player}(s_{t+1}, a_{t+1}) - Q_{player}(s_t, a_t). \quad (5)$$

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. Data and codes are available at [https://github.com/Ning-D/Evaluate\\_Badminton\\_Stroke](https://github.com/Ning-D/Evaluate_Badminton_Stroke).

### 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 pose) achieves the best performance, indicating that all the input components contribute in estimating an accurate Q-function.

**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	– Player position	– Player pose	Full
$\mathcal{L}$	0.0289	0.0436	0.0506	<b>0.0261</b>

Next, we examined the effect of the stroke type on our model. As the action is a primary element in reinforcement learning, we used a fully supervised LSTM model (non-reinforcement learning) to examine the effect of action (stroke types) by eliminating it from the total input. The results show that the stroke type feature can help in improving the accuracy of the fully supervised LSTM model (loss function value: 0.0995) when compared to a model that does not consider this feature (0.1274).

**C. CHARACTERISTICS OF THE PLAYERS IDENTIFIED USING OUR METHOD**

The matches considered here include one final, two semi-finals, 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).

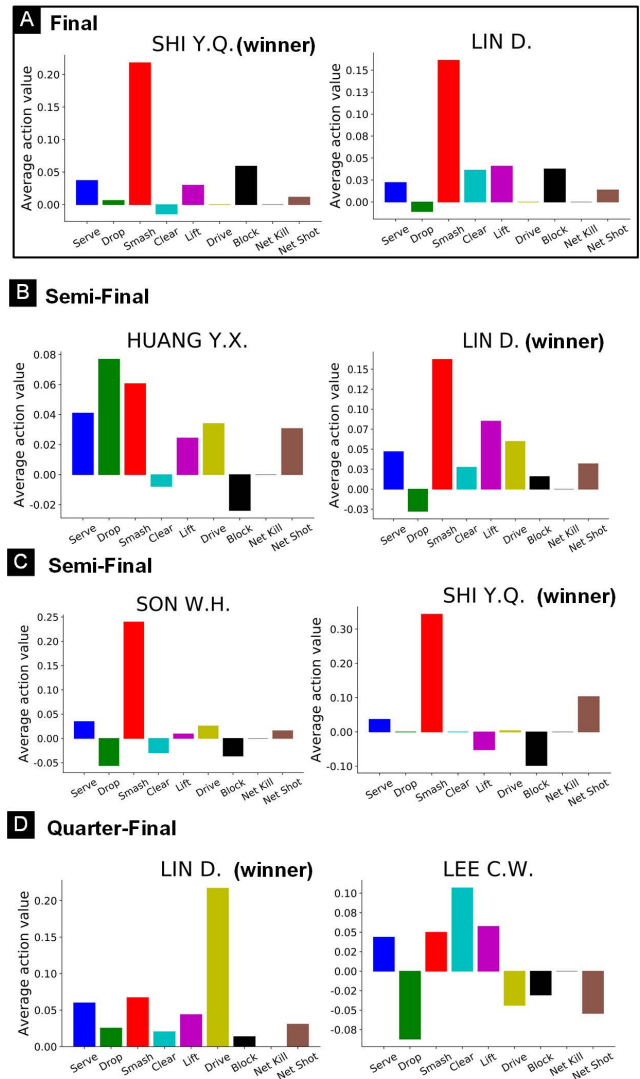
For example, we identified that, when Lin was playing against Huang and Lee (Figs. 5 (B) and (D)), the strokes performed by Lin were superior to those by his opponents. Lin had fewer stroke types whose action value were substantially below zero, especially when he was playing against Lee. His offensive 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 as 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  $\rho = 0.150$  ( $p > 0.05$ ), 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



**FIGURE 5. Average action value of each stroke types in four badminton finals: (A) Final, (B) Semi-final, (C) Semi-final, (D) Quarter-final.**

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.

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

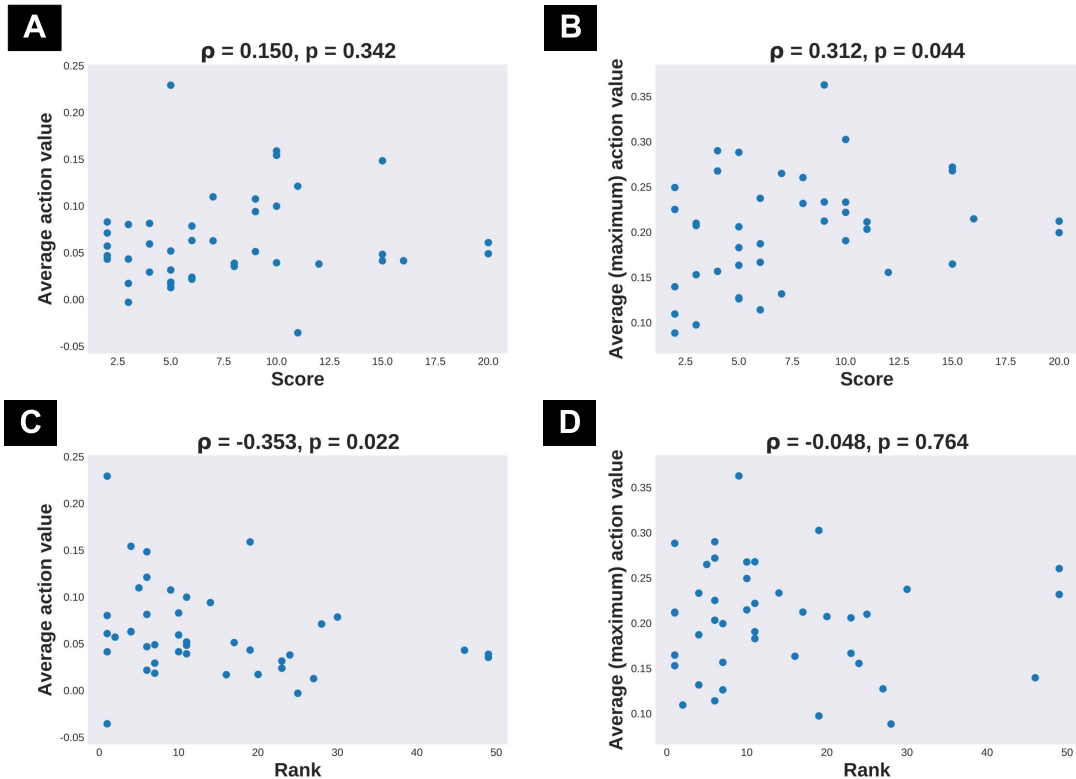


FIGURE 6. Upper (A, B) Correlation between the score and average action value/average (maximum) action value. Lower (C, D) Correlation between the rank of the player and average action value/average (maximum) action value.

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 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.

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

Model	— Player pose (baseline)		Full model	
	Score	Rank	Score	Rank
Average action value	-0.154	-0.122	0.150	<b>-0.353</b>
Average (maximum) action value	-0.061	0.163	<b>0.312</b>	-0.048

We summarized 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. CONCLUSION

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. We shared a badminton dataset with ground truth annotations of both stroke types and game scores publicly.

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

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

## REFERENCES

- [1] J. Ensum, R. Pollard, and S. Taylor, "Applications of logistic regression to shots at goal at association football," in *Proc. 5th World Congr. Sci. Football*, 2005, p. 214.
- [2] L. Shen, H. Zhang, M. Zhu, J. Zheng, and Y. Ren, "Measurement and performance evaluation of lob technique using aerodynamic model in badminton matches," in *Proc. Int. Symp. Comput. Sci. Sport*. Cham, Switzerland: Springer, 2019, pp. 53–58.
- [3] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.
- [4] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, and G. Van Den Driessche, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.
- [5] A. E. Sallab, M. Abdou, E. Perot, and S. Yogamani, "Deep reinforcement learning framework for autonomous driving," *Electron. Imag.*, vol. 2017, no. 19, pp. 70–76, 2017.
- [6] J. Liu, B. Logan, N. Liu, Z. Xu, J. Tang, and Y. Wang, "Deep reinforcement learning for dynamic treatment regimes on medical registry data," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Aug. 2017, pp. 380–385.
- [7] S. Gu, E. Holly, T. Lillicrap, and S. Levine, "Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2017, pp. 3389–3396.
- [8] M. Hausknecht and P. Stone, "Deep recurrent Q-learning for partially observable MDPs," in *Proc. AAAI Fall Symp. Ser.*, 2015, pp. 1–9.
- [9] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, and A. Graves, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [10] G. Liu and O. Schulte, "Deep reinforcement learning in ice hockey for context-aware player evaluation," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 3442–3448, doi: 10.24963/IJCAI.2018/478.
- [11] T. Gerstenberg, T. Ullman, M. Kleiman-Weiner, D. Lagnado, and J. Tenenbaum, "Wins above replacement: Responsibility attributions as counterfactual replacements," in *Proc. Annu. Meeting Cogn. Sci. Soc.*, 2014, vol. 36, no. 36, pp. 1–7.
- [12] D. L. Cerynik, G. E. Lewullis, B. C. Joves, M. P. Palmer, and J. A. Tom, "Outcomes of microfracture in professional basketball players," *Knee Surg., Sports Traumatol., Arthroscopy*, vol. 17, no. 9, pp. 1135–1139, Sep. 2009.
- [13] D. Cervone, A. D'Amour, L. Bornn, and K. Goldsberry, "POINTWISE: Predicting points and valuing decisions in real time with NBA optical tracking data," in *Proc. 8th MIT Sloan Sports Anal. Conf.*, Boston, MA, USA, vol. 28, 2014, pp. 1–9.
- [14] T. Decroos, L. Bransen, J. Van Haaren, and J. Davis, "Actions speak louder than goals: Valuing player actions in soccer," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2019, pp. 1851–1861.
- [15] K. Toda, M. Teranishi, K. Kushiro, and K. Fujii, "Evaluation of soccer team defense based on prediction models of ball recovery and being attacked: A pilot study," *PLoS ONE*, vol. 17, no. 1, Jan. 2022, Art. no. e0263051.
- [16] K. Fujii, "Data-driven analysis for understanding team sports behaviors," *J. Robot. Mechatron.*, vol. 33, no. 3, pp. 505–514, Jun. 2021.
- [17] G. Liu, Y. Luo, O. Schulte, and T. Kharrat, "Deep soccer analytics: Learning an action-value function for evaluating soccer players," *Data Mining Knowl. Discovery*, vol. 34, no. 5, pp. 1531–1559, Sep. 2020.
- [18] X. Sun, J. Davis, O. Schulte, and G. Liu, "Cracking the black box: Distilling deep sports analytics," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 1–9.
- [19] D. Cervone, A. D'Amour, L. Bornn, and K. Goldsberry, "A multiresolution stochastic process model for predicting basketball possession outcomes," *J. Amer. Stat. Assoc.*, vol. 111, no. 514, pp. 585–599, Apr. 2016.
- [20] J. Fernández, L. Bornn, and D. Cervone, "Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer," in *Proc. 13th MIT Sloan Sports Anal. Conf.*, 2019, pp. 1–20.
- [21] D. Cervone, L. Bornn, and K. Goldsberry, "NBA court reality," in *Proc. 10th MIT Sloan Sports Anal. Conf.*, 2016, pp. 1–8.
- [22] J. Fernandez and L. Bornn, "Wide open spaces: A statistical technique for measuring space creation in professional soccer," in *Proc. MIT Sloan Sports Anal. Conf.*, 2018, pp. 1–19.
- [23] T. Taki and J. Hasegawa, "Visualization of dominant region in team games and its application to teamwork analysis," in *Proc. Comput. Graph. Int.*, Jun. 2000, pp. 227–235.
- [24] J. Wang, I. Fox, J. Skaza, N. Linck, S. Singh, and J. Wiens, "The advantage of doubling: A deep reinforcement learning approach to studying the double team in the NBA," 2018, *arXiv:1803.02940*.
- [25] R. Voeikov, N. Falaleev, and R. Baikulov, "TTNet: Real-time temporal and spatial video analysis of table tennis," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 884–885.
- [26] N. A. Rahmad, N. A. J. Sufri, N. H. Muzamil, and M. A. As'ari, "Badminton player detection using faster region convolutional neural network," *Indonesian J. Elect. Eng. Comput. Sci.*, vol. 14, no. 3, pp. 1330–1335, 2019.
- [27] F. Yan, W. Christmas, and J. Kittler, "A tennis ball tracking algorithm for automatic annotation of tennis match," in *Proc. Brit. Mach. Vis. Conf.*, vol. 2. Guildford, U.K.: Univ. Surrey, 2005, pp. 619–628.
- [28] H. Shah, P. Chokalingam, B. Paluri, N. Pradeep, and B. Raman, "Automated stroke classification in tennis," in *Proc. Int. Conf. Image Anal. Recognit.* Berlin, Germany: Springer, 2007, pp. 1128–1137.
- [29] K. M. Kulkarni and S. Shenoy, "Table tennis stroke recognition using two-dimensional human pose estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2021, pp. 4576–4584.
- [30] W.-T. Chu and S. Situmeang, "Badminton video analysis based on spatiotemporal and stroke features," in *Proc. ACM Int. Conf. Multimedia Retr.*, Jun. 2017, pp. 448–451.



- [31] T. Yoshikawa, K. Tsutsui, K. Takeda, and K. Fujii, "Extraction of swing motion contributing to prediction of shuttle drop position in badminton," in *Proc. 30th Int. Joint Conf. Artif. Intell. (IJCAI) Workshop AI Sports Anal. (AISA)*, 2021, pp. 1–2.
- [32] T. Barnett, A. Brown, and S. Clarke, "Developing a model that reflects outcomes of tennis matches," in *Proc. 8th Australas. Conf. Math. Comput. Sport*, Coolangatta, QLD, Australia, 2006, pp. 3–5.
- [33] M. Sharma, N. Kumar, and P. Kumar, "Badminton match outcome prediction model using Naïve Bayes and feature weighting technique," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 8, pp. 8441–8455, 2021.
- [34] L. Vergauwen, A. J. Spaepen, J. Lefevre, and P. Hespel, "Evaluation of stroke performance in tennis," *Med. Sci. Sports Exerc.*, vol. 30, no. 8, pp. 1281–1288, Aug. 1998.
- [35] M. Pfeiffer, H. Zhang, and A. Hohmann, "A Markov chain model of elite table tennis competition," *Int. J. Sports Sci. Coaching*, vol. 5, no. 2, pp. 205–222, Jun. 2010.
- [36] T. McGarry and I. M. Franks, "Development, application, and limitation of a stochastic Markov model in explaining championship squash performance," *Res. Quart. Exerc. Sport*, vol. 67, no. 4, pp. 406–415, 1996.
- [37] J. Wang, K. Zhao, D. Deng, A. Cao, X. Xie, Z. Zhou, H. Zhang, and Y. Wu, "Tac-Simur: Tactic-based simulative visual analytics of table tennis," *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 1, pp. 407–417, Jan. 2020.
- [38] S. Wenninger and M. Lames, "Markov simulation by numerical derivation in table tennis," in *Proc. 10th Int. Symp. Comput. Sci. Sports (ISCSS)*. Cham, Switzerland: Springer, 2016, pp. 161–169.
- [39] J. Wang, D. Deng, X. Xie, X. Shu, Y.-X. Huang, L.-W. Cai, H. Zhang, M.-L. Zhang, Z.-H. Zhou, and Y. Wu, "Tac-Valuer: Knowledge-based stroke evaluation in table tennis," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining*, Aug. 2021, pp. 3688–3696.
- [40] W.-Y. Wang, T.-F. Chan, H.-K. Yang, C.-C. Wang, Y.-C. Fan, and W.-C. Peng, "Exploring the long short-term dependencies to infer shot influence in badminton matches," 2021, *arXiv:2109.06431*.
- [41] H.-S. Fang, S. Xie, Y.-W. Tai, and C. Lu, "RMPE: Regional multi-person pose estimation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2334–2343.
- [42] Y.-C. Huang, I.-N. Liao, C.-H. Chen, T.-U. Ik, and W.-C. Peng, "TrackNet: A deep learning network for tracking high-speed and tiny objects in sports applications," in *Proc. 16th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Sep. 2019, pp. 1–8.
- [43] K. Routley and O. Schulte, "A Markov game model for valuing player actions in ice hockey," in *Proc. 31st Conf. Uncertainty Artif. Intell.*, 2015, pp. 782–791.



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



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

His research interest includes behavior signal processing—driving behavior.



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

behavioral sciences, and sports sciences.

• • •