2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC)

Badminton Action Analysis Using LSTM

1st Abhay Malik
On My Own Technology
Mumbai, India
abhay.x.malik@gmail.com

2nd Reetu Jain
Chief Mentor
On My Own Technology
Mumbai, India
reetu.jain@onmyowntechnology.com

2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC) | 979-8-3503-2379-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICAISC58445.2023.10201000

Abstract—**Badminton is a popular racquet sport that involves quick movements, agility, and reflexes. A badminton action analysis is a process in which quantify the various actions performed by the player using various sensors or image processing-based approaches. In this paper, a sensor-based approach to collect player action data for various shots, analyse the same and create a Long short-term memory-based model to classify various shots. The trained Long short-term memory model has a Micro F1 score of 0.84. The Model would then generate a shot accuracy score between 1-10. A professional badminton coach also rated the shots, and a comparison was made between the score given by the coach and the ML model. Additionally, arm section-wise comparison of the shot was made. This allows for a more granular analysis of the shot. A comparison of the arm-wise shot is also conducted. It can be seen from the results that the scores given by the coach and the Long short-term memory-based ML model matched by 98.6%. So it can be concluded that such a sensor-based wearable badminton teaching technique can be deployed to enable underprivileged beginner badminton players to achieve higher shot accuracy by allowing them to correct postures during the swing action of the arm.**

Keywords—Badminton, Action Analysis, LSTM, Sports Analysis, Machine Learning, Neural Network.

I. Introduction

Millions of people worldwide actively participate in badminton, a very well-liked and energetic sport that has grown in popularity recently. Players of various ages and skill levels find the sport demanding and interesting since it calls for a special fusion of physical and mental abilities. The variety of gaming forms in badminton is one of its most noteworthy features. The most popular formats are singles and doubles, and they each call for a unique set of abilities and tactics. Players in singles must cover the entire court while outmanoeuvring their opponents using superb footwork, positioning, and agility. Conversely, in doubles, players must work together and communicate properly to cover the court more efficiently and provide more opportunities for attacking strokes [7]. There is a long history of competition in badminton, and many professional players participate in leagues and tournaments around the world [11]. Millions of spectators travel from all over the world to watch these competitions, which highlight the talent and athleticism of the best players in the world [9]. Badminton is a sport that can be played both competitively and recreationally, making it a great form of exercise for those who want to stay active. Machine learning methods have drawn more attention in recent years for the analysis of sports data, including badminton. Using Long Short-Term Memory (LSTM) [8], networks are one method for studying badminton data. LSTM is a type of recurrent neural network (RNN) that can learn long-term dependencies in sequential data. By using LSTMs, it is possible to analyse the sequences of actions taken by players during a badminton game and identify patterns that may be useful for coaches and players to improve their performance.

Badminton action analysis using LSTMs involves collecting and preprocessing data from badminton matches, such as player positions, shuttlecock trajectories, and other relevant information. The trained machine learning model would enable the player to self-analyze a particular shot, thus allowing them to improve by correcting the mistakes in posture and the overall swing action. This would allow underprivileged players to access high-quality coaching without requiring an actual coach. Additionally, using readily available sensors and controller makes the wearable setup cheap. Since a Neural network-based model with just six layers is trained, it requires only 9 distinct features for the entire shot analysis, thus reducing the complexity of the model. This also allows the ML model to be run on a mobile computing device.

II. Literature Review

A study conducted by Chiu et al. [1] found that the most important badminton skill is the capacity for acceleration and deacceleration, as well as the capacity for changing directions on the court. The objective of their study was to determine whether a smartphone-based assessment system for badminton six-point footwork would be practical. The study was conducted on thirty badminton players. Based on the mean and maximum acceleration data, the results demonstrated that performance between the quicker and slower groups could be distinguished by using a smartphone's built-in accelerometer. The results showed that the mean acceleration of players in the fast group was considerably higher than that of players in the slow group, with effect sizes ranging from 0.75 to 1.70. Wang et al. [2] proposed ShuttleNet, a novel framework for stroke forecasting in badminton that combines information about the rally and player using modified encoder-decoder extractors and a fusion network. ShuttleNet outperformed other approaches in tests using a badminton dataset, demonstrating its effectiveness in predicting match outcomes. This study highlights the importance of stroke forecasting in sports and offers a promising approach for objective decision-making in turn-based sports. The proposed badminton environment by Wei-Yao et al. [3] is a secure and repeatable simulator that can test and improve player performance and fan engagement. By modelling rallies from multiple points of view and defining states, actions, and training schedules, the environment allows for efficient assessment of original algorithms and mimics earlier matches for method investigation. The use of cross-entropy as a measure of uncertainty helps in evaluating shot-type prediction. This environment and methodology can benefit coaches, players, and researchers alike in improving the sport of badminton. To test and enhance badminton performance, Huang, Li-Chun, and colleagues [4] created a badminton environment that mimics rallies from various angles and generates states, actions, and training strategies to test in a trustworthy and repeatable simulator. They also introduce a number of agents and a diagram of the reinforcement learning environment with two supplementary views to record match data. This adaptable training environment is a crucial addition to sports evaluation since it can handle various scenarios and train several agents simultaneously.

Studies conducted by Zhesen Chu et al. show that [5] Biomechanics is important in badminton, and a little research over the years has improved our knowledge of the sport. The following review summarises our current understanding of the four common badminton strokes (forehand serve, power stroke, forehand and backhand overhead strokes, and general endurance and fitness). The research conducted by Sorensen et al. [6] aims to use inertial sensors to generate a 3D scene and control a 3D human body model using a modified deep LSTM network. They use a single inertial sensor attached to a badminton racket for data collection and a window segmentation approach to identify six common swing movements with an accuracy of over 90%. Additionally, the study examines biomechanical differences in stroke technique among young badminton players, focusing on longitudinal axis rotations and proximal-distal sequencing. Skilled players achieved higher angular velocities in the backhand stroke. Skilled and less skilled players used a proximal-distal sequence in the forehand clear stroke, resulting in joint reaction forces that transported energy from the proximal to the distal segments.

The study carried out by Wei Yao Wang et al. [10] shows that in analysing a player's performance in a badminton match, it's critical to recognise keystrokes in a rally. Analysing badminton data has not been studied as much as studies that have measured player performance in other sports. By framing the issue as the prediction of the outcome of a rally, we propose a deep learning model in this paper that consists of a novel short-term extractor and a long-term encoder for capturing a shot-by-shot sequence in a badminton rally. In the study by Li-Chun Huang et al. [12] diverse attempts to improve player performance and fan engagement have been spawned by new techniques for effectively monitoring sports. However, because testing in real-time matches is prohibitively expensive and impracticable, existing approaches can only evaluate offline performance. To test in a safe and reliable simulator, we focus on turn-based sports and provide a badminton environment by simulating rallies from multiple points of view and building the states, actions, and training techniques. This benefits coaches, players, and academics who may swiftly evaluate their special algorithms by simulating prior matches for technique inspection. The research of Wei-Yao Wang and associates [13] numerous instances of fruitful research, including health status monitoring and result prediction, have been inspired by the growing need to assess sports-related findings. The issue of choosing what and where to return strokes objectively in turn-based sports is still up for debate. However, they are unable to model data based on badminton-specific properties. Existing research can address the issue by defining stroke forecasting as a sequence prediction job. We provide ShuttleNet, a position-aware framework that combines player styles with rally progress, to get around these restrictions. It combines player data with rally progress using two modified encoder-decoder extractors. According to a study done by Chao Ma et al. [14], people who participate in different sports are becoming more numerous every day due to the quick development of sports. Badminton has become one of their most popular sports since it has fewer field restrictions and is easy to learn. This work develops an accurate badminton movement recognition algorithm for wearable sports activity classification. A single acceleration sensor is attached to the end of the badminton racket handle to capture data on the game's activity. The sliding window segmentation method is used to extract the striking signal. In the study carried out by Lei Fang et al. [15], the application of intelligent identification technology in sports has generated considerable attention and has seen significant advancements in the transportation industry. This work's major focus is the examination of human motion detection systems in recordings of sports and dance performances. An angle-adaptive and continuous-scale space template matching technique in geometric, algebraic space is developed using instance templates, a unique cfrdF approach, and CSS-based similarity of human body traits as features. This method calculates the value of similarity between horizontal plates and detected photographs before establishing a threshold to match the region where the human body is positioned.

III. Proposed Methodology

To analyse badminton actions like smash, drop etc., it is important to monitor the player's body and the racket movement. There are various available options that involve using a camera to capture the video of the player and then using various landmark detection models, which can later be analysed. Other options involve using a sensor module attached to various locations on the player's body to collect data. Mostly the sensors would be small MEMS-based accelerometers. This paper explores a variation of the second as it gives direct data about acceleration. Landmark data given by the video could also be used to calculate the acceleration of various body joints, but this adds another layer of complexity, and the accuracy of the methods will depend on the accuracy of the landmark detection model.

To collect the acceleration data of the various actions done during a badminton match, small units consisting of IMU (inertial measurement units) were created. There are 3 separate modules designed. The first one is attached to the Badminton shaft, the second one is attached to the wrist of the player, and the third one is attached to the upper arm of the player. The units attached to the wrist and the upper arm are similar. The unit attached to the badminton racket shaft is designed differently by taking into consideration the weight, which might affect the performance of the player. Figure 1 shows the placement of the sensor units on the badminton player's body and the racquet. These locations are specifically used as they allow the sensors to capture acceleration on all the separate joints including the upper arm, forearm and wrist.

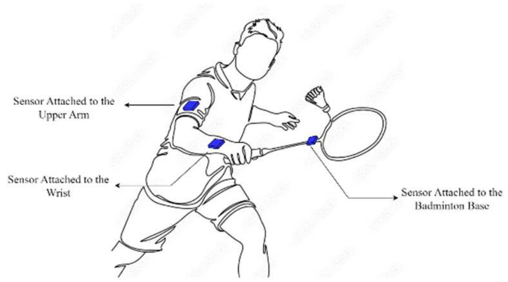

Fig. 1. Location of the 3 units mounted on the player

M5-stamp pico microcontrollers were used with MPU6050 IMU sensors to collect the acceleration data. The Circuits used in each unit were the same. The unit attached to the wrist and the arm was made of several components. The microcontroller and the IMU were in a separate enclosure, and the battery and the switch were packaged separately. The unit connected to the badminton shaft was designed as a single enclosure. Figure 2 shows the exploded view of the units which attach to the badminton racket.

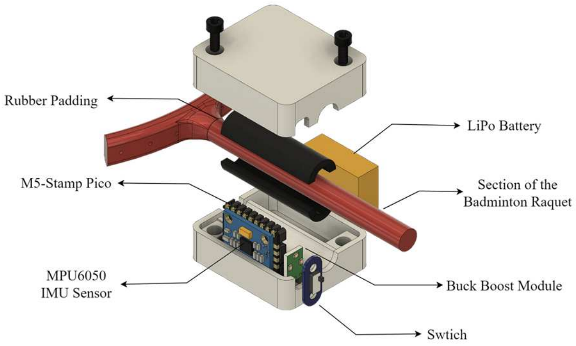

Fig. 2. 3D model of the sub-components in the Wrist and the upper arm assembly

The wrist and the upper arm unit are split into 3 separate subunits. This allows the modules to be compact and comfortable for the user to wear on their arm while playing. The first subunit encloses the battery and the charge-discharge module; the second unit encloses the switch, the buck-boost converter and the charging JST connector. The third subunit has the controller and the IMU sensor. All the subunits have a sewing attachment which allows them to be stitched to any readily available wrist or armband. Splitting individual components allows the subunit to be compact and can be placed independently on any location on the wrist. Figure 3 shows the exploded view of the wrist and the upper arm units. The units are divided into multiple smaller parts to reduce the inference to the player's arm movement. Figure 4 shows the units attached to the badminton and the armband. This is the configuration that was used for data collection.

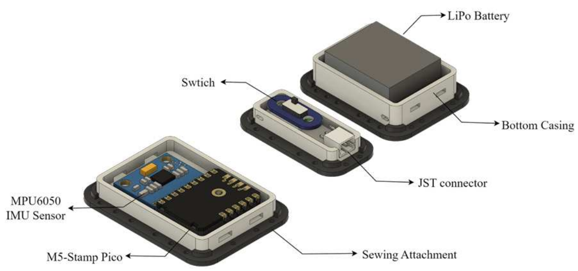

Fig. 3. 3D model of the sub-components in the Wrist and the upper arm assembly

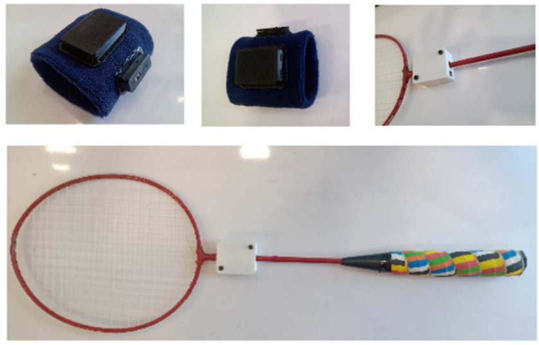

Fig. 4. Actual Prototype Images

A. Data acquisition

Data was collected by attaching the units to badminton players. The players were asked to perform a single action repeatedly. A total of 10 players' data were collected. All the players selected for data collection were professional players. This is mandatory as ML model needs to train on the best data available for each shot. The dataset consisted of data for clear, drive, drop and smash shots. Each one of these shots was performed by the player for a total of 50 repetitions. The table I shows the various shots and the total repetitions of each recorded shot.

TABLE I. DIFFERENT TYPES AND THE COUNT OF VALID SHOTS RECORDED

Shots Performed	Number of samples collected
Clear	500
Drive	480
Drop	500
Smash	450

All the individual units connect to the local hotspot created on the laptop. The laptop also runs the Python code, which collects the data from the individual units. The data is transferred using the TCP protocol. Individual units act as a TCP server. Python code running on the laptop runs multiple threads, each acting as an individual TCP client connecting to the sensor units. Multithreading allows the data to be collected simultaneously from all the sensor units. The collected data for individual units were stored as separate .npy files (numpy binary file format). All the controllers use NTP (network time protocol) to sync the time stamp of the data collected. Since determining the start and stop of a particular action from the continuous stream of data would be very difficult so, during data collection, a person would manually trigger the start and stop timing for each action. The NumPy file consists of data for all 50 repetitions, one after the other, in the same file. Since the shots performed vary in time, precautions were taken to monitor individual shots. A 3-second window was allocated for each type of shot. If the shot were completed before this 3-second mark, the remaining rows were filled with zeros during the pre-processing step.

Each sensor unit samples the IMU at 50hz and streams the data over TCP. Since a 3-second window was allocated for each action, which corresponds to150 data points for each action. But since the action performed by different individuals would have different lengths, pre-processing steps were used to even out the length of the various actions. If any particular action would not have 150 data points, remaining section were filled with zeros denoting that no acceleration was being exerted on the sensors. Figure 5 shows a preprocessing step that was used to sample the collected data so that all different shots had a uniform number of data points.

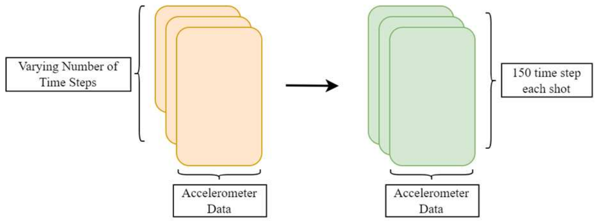

Fig. 5. Preprocessing is used to even out the number of timesteps in each action/shot

Pre-processing steps to even out the length of the data points are mandatory, as the LSTM model requires the length of the data to be consistent. The data collected from different individuals were combined into a single numpy file. 4 separate numpy files were created, one each for the clear, drive, drop and smash shots. Each action data collected consisted of 9 features. The IMU sensor gives 3 features corresponding to the x,y and z acceleration. The acceleration components also depend on the mounting position of the sensors. Hence, the sensors were mounted in the same orientation on the badminton shaft, wrist and upper arm. Each of the x,y and z acceleration data is independent of each other and hence acts as a separate feature.

The numpy files for each type of shot were converted into a three-dimensional NumPy array compatible with the LSTM model's required input shape (window, frames, features). This allows the array to be used directly with the model. Figure 6 represents the conversion of the 2D data array into a 3D array which is required by the LSTM model for training.

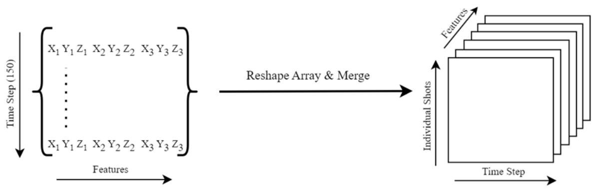

Fig. 6. Reshaping dataset for the LSTM model

The number of rows in the NumPy array denotes the number of times the particular action was recorded. Every row is composed of 150 individual timesteps, and each timestep is composed of 9 accelerometer features.

B. Machine learning model

The model was trained using Long Short-Term Memory (LSTM). LSTM stands for Long Short-Term Memory, and it is a type of recurrent neural network (RNN) that is designed to address the issue of the vanishing gradient problem in traditional RNNs. LSTM networks are designed to remember information for longer periods of time, making them particularly effective for tasks that involve sequential data, such as speech recognition, natural language processing, and time series analysis. To meet LSTM's input shape requirements, our data was organised into an array of (timesteps, features), with 150 timesteps and 9 features per timestep. Similar to the Recurrent Neural Network (RNN) model, LSTM can retain information from previous timesteps to learn complex actions that span multiple timesteps in the recorded data. Additionally, since neural networks cannot train on text input, the label names of each action were converted into categorical data by using one hot encoding.

To implement the LSTM model, the a Keras sequential model was used. The first layer of the model was an LSTM layer with an output space of 40. To prevent overfitting of the data, a dropout layer with a 30% dropout rate was added after the LSTM layer. Following the dropout layer, 2 dense layers with an output dimensionality of 40 were added. The dense layer is a deeply connected neural network layer. Another dropout layer with a 30% dropout rate was added after the dense layer, which was connected to the final Dense layer. The final Dense layer had 4 output dimensionality, corresponding to the number of shots being trained. Figure 7 shows the LSTM model architecture used for training.

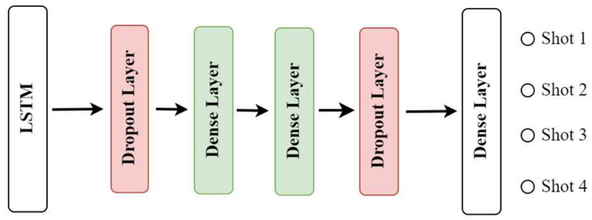

Fig. 7. LSTM model used for the training

After preparing the Sequential model; it was trained on the training dataset using Adam's optimiser. The training data set was shuffled before training to ensure that the model didn't learn the relationship between different shot data. The input data containing dance step labels were converted into categorical form as the LSTM model requires binary data for better learning. Figure 8 shows the confusion matrix for the shot classification accuracy of the model.

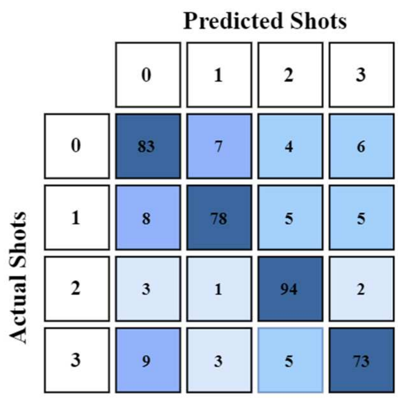

Fig. 8. Confusion matrix for the LSTM model

Once the model achieved acceptable accuracy, a standard scaling-based approach was used to convert the confidence level of LSTM classification into scores ranging from 1-10. The minimum and maximum values of the confidence score obtained during testing were used to scale the confidence score and obtain the desired scoring.

Additionally, 3 separate ML models were trained for each section of the arm to check the accuracy of the swinging motions performed by that particular section of the arm. There are 3 features in each time step which is used to train the model. The data for training this model is derived from the full dataset. The LSTM model and the training parameters remain the same. The secondary model (for individual arm sections) is used in conjunction with the primary model. Once the primary model predicts the shot type, the second model detects the shot accuracy for each individual section of the arm. So the final result from a single shot involves overall shot accuracy and individual arm section accuracy. Thus, allowing for a more granular shot analysis.

TABLE II. COMPARISON OF SCORINGS DONE BY A PROFESSIONAL BADMINTON PLAYER VS THE SCORING DONE BY THE ML MODEL

	Shot 1							
	Overall		***Arm1***		***Arm2***		***Arm3***	
	Coach	***ML model***	***Coach***	***ML model***	***Coach***	***ML model***	***Coach***	***ML model***
Individual - 1	6	6	6	6	6	6	6	6
	8.3214	8.3214	8.3214	8.3214	8.3214	8.3214	8.3214	8.3214
	5	5	5	5	5	5	5	5
Individual - 2	6.35729	6.35729	6.35729	6.35729	6.35729	6.35729	6.35729	6.35729
	8	8	8	8	8	8	8	8
	8.1240	8.1240	8.1240	8.1240	8.1240	8.1240	8.1240	8.1240

IV. TEST RESULTS

To test the model, data was collected from 1 beginner-level badminton player and 1 experienced badminton player while monitored by a professional badminton coach. Each player performed a single shot for 2 repetitions while wearing the device. A simple standard scaler-based approach was used to scale the prediction accuracy score between 1-10. The coach monitoring the action was asked to give an overall shot accuracy score between 1 and 10 for each action performed. Additionally, the coach also gave scores for 3 individual sections of the arm. The above table II shows the results obtained from the analysis for a 3-shot performed by 2 separate individuals. In the above table, Arm 1 denotes the badminton and wrist motion, Arm 2 denotes the forearm motion, and Arm 3 denotes the upper arm motion. Individual 1 was a beginner-level player, and Individual 2 was an experienced player.

The LSTM-based model had a Micro F1 score of 0.84. This only represents the shot classification accuracy of the model. To understand the shot analysis aspect of the model better an analysis of the sores given by the model and the coach was done. This would allow us to better understand if such a system can traditional methods of having a coach to improve skills. The overall score given by the coach and the ML model matched by almost 98.6% which shows that the suggestions given by the model are very similar to the suggestions given by a coach. For individual sections of the arm, the scores matched by 94.78%, 96.02% and 92.85% respectively for the wrist, forearm and upper arm sections respectively.

V. CONCLUSION

Comparing the test results shows that the model's prediction and the score for the overall accuracy of the shot are very similar to the scores given by a professional coach. Generally, the shots were given a low score by the coach; the ML model also gave a low score. For high-scoring shots, the ML model generally would give a higher score than the coach. This might be because the ML model only looks at the acceleration values of the swing. The ML model can detect minute changes occurring even in a quick shot allowing the model to give a more accurate result than would be possible by looking at the shot with the bare eyes. Thus the proposed system can help underprivileged beginner badminton players to improve their shots without the need to rely on expensive coaching. The sensor setup used to collect data can be created by using cheap sensor modules, thus reducing the cost of the overall wearable setup.

The proposed technique allows us to detect one of the 4 shots being performed and be able to check the model's prediction confidence for a new shot that is performed. More number of available shots and variations can be included while expanding on the data used to train the model. One of the prospects is to be able to detect which segment of the hand is deviating from the ideal conditions and be able to suggest corrective actions. The posture of the body during a shot affects the impact of the shot. Being able to capture and add features representing the rotational motion of the body to the training dataset. This would allow the model to give comprehensive analysis and suggestions to improve the shot. Another important aspect of badminton is the footwork associated with each shot which is the most difficult to master. The current proposed system does not take into account the foot positioning of the player. In one of the future iterations of the proposed system, additional sensors can be incorporate to track the foot movement of the player thus giving a more comprehensive analysis to the players.

ACKNOWLEDGMENT

I would like to express my gratitude to the mentors of On My Own Technology Pvt. Ltd. for extending their help in carrying out this project.

REFERENCES

[1] Chiu, Ya-Lan, Tsai, Chia-Liang, Sung, Wen-Hsu and Tsai, Yi-Ju. (2020). Feasibility of Smartphone-Based Badminton Footwork Performance Assessment System. Sensors. 20. 1-11. 10.3390/s20216035.

[2] Wang, W., Chan, T., Yang, H., Wang, C., Fan, Y., and Peng, W. (2021c). Exploring the Long Short-Term Dependencies to Infer Shot Influence in Badminton Matches. ArXiv (Cornell University). https://doi.org/10.48550/arxiv.2109.06431

[3] Wang, Wei-Yao, Shuai, Hong-Han, Chang, Kai-Shiang, Peng and Wen-Chih. (2021). ShuttleNet: Position-aware Fusion of Rally Progress and Player Styles for Stroke Forecasting in Badminton.

[4] Huang, Li-Chun, Hseuh, Nai-Zen, Chien, Yen-Che, Wang, Wei-Yao, Wang, Kuang-Da, Peng and Wen-Chih. (2022). A Reinforcement Learning Badminton Environment for Simulating Player Tactics (Student Abstract). 10.48550/arXiv.2211.12234.

[5] Zhesen Chu, Min Li, "Image Recognition of Badminton Swing Motion Based on Single Inertial Sensor", Journal of Sensors, vol. 2021, Article ID 3736923, 12 pages, 2021. https://doi.org/10.1155/2021/3736923

[6] Sorensen, K., de Zee, M., & Rasmussen, J. (2011). A biomechanical analysis of clear strokes in badminton executed by youth players of different skill levels. Paper presented at Congress of the International Society of Biomechanics, ISB, Brussels, Belgium.

[7] Valldecabres, R., Casal, C. A., Chiminazzo, J. G., and de Benito, A. M. (2020). Players' On-Court Movements and Contextual Variables in Badminton World Championship. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.01567

[8] Chu, Z., and Li, M. (2021, September 28). Image recognition of badminton swing motion based on single Inertial Sensor. Journal of Sensors. https://www.hindawi.com/journals/js/2021/3736923/

[9] Manan V, Manit A, Digpal R. Biomechanics in Badminton- A Review. Orthop and Spo Med Op Acc J 2(1)- 2018. OSMOAJ.MS.ID.000129. DOI: 10.32474/OSMOAJ.2018.02.000129.

[10] 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 IEEE International Conference on Data Mining (ICDM), Auckland, New Zealand, 2021, pp. 1397-1402, doi: 10.1109/ICDM51629.2021.00178.

[11] Chiu, Y.L., Tsai, C.L., Sung, W.H, Tsai, Y.J. Feasibility of Smartphone-Based Badminton Footwork Performance Assessment System. Sensors 2020, 20, 6035. https://doi.org/10.3390/s20216035

[12] Huang, L., Hseuh, N., Chien, Y., Wang, W., Wang, K., and Peng, W. (2022). A Reinforcement Learning Badminton Environment for Simulating Player Tactics (Student Abstract). ArXiv. /abs/2211.12234

[13] Wang, W., Shuai, H., Chang, K., and Peng, W. (2021). ShuttleNet: Position-aware Fusion of Rally Progress and Player Styles for Stroke Forecasting in Badminton. ArXiv. /abs/2112.01044

[14] Chao Ma, Dayang Yu, Hao Feng, "Recognition of Badminton Shot Action Based on the Improved Hidden Markov Model", Journal of Healthcare Engineering, vol. 2021, Article ID 7892902, 8 pages, 2021. https://doi.org/10.1155/2021/7892902

[15] Fang, L., and Sun, M. (2021). Motion recognition technology of badminton players in sports video images. Future Generation Computer Systems, 124, 381-389. https://doi.org/10.1016/j.future.2021.05.036