Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal **Vol. 7, No. 1, January 2024, pp: 27-37; eISSN: 2656-8632**

RESEARCH ARTICLE OPEN ACCESS

Manuscript received June 21, 2024; revised October 15, 2024; accepted October 16, 2024; date of publication November 11, 2024
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v7i1.488

How to cite: Bayu Erfianto, Bedy Purnama , and Ilo Raditio Wirawan, "Time Series Classification of Badminton Pose Using Long Short-Term Memory with Landmark Tracking", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 7, no. 1, pp. 27-37, January 2024.

Time Series Classification of Badminton Pose Using Long Short-Term Memory with Landmark Tracking

Bayu Erfianto, Bedy Purnama, and Ilo Raditio Wirawan

School of Computing, Telkom University Bandung, Bandung, Indonesia

Corresponding author: Bedy Purnama (e-mail: bedypurnama@telkomuniversity.ac.id).

The author would like to express gratitude to PPM Telkom University. This research was funded by Telkom University's Internal Basic Research, Period 2, 2023.

ABSTRACT Traditional methods of analyzing badminton matches, such as video movement analysis, are time-consuming, prone to errors, and rely heavily on manual annotation. This creates challenges in accurately and efficiently classifying badminton actions and player poses. This paper aims to develop an accurate time series classification method for badminton poses using landmark tracking. The proposed method integrates Long Short-Term Memory (LSTM) networks with landmark tracking to classify badminton poses in a time series, addressing the limitations of traditional video analysis techniques. The dataset consists of 30 respondents performing three distinct activities—lob, smash, and serve—under two conditions: good and bad execution. The approach combines LSTM networks with landmark tracking data, utilizing intra-class variation from a multi-view dataset to enhance pose classification accuracy. The LSTM model achieved high accuracy in classifying badminton poses, successfully detecting serves, lobs, and smashes in real-time with over 90% accuracy. Additionally, the system improved match analysis, achieving 85% accuracy in detection and classification, demonstrating the effectiveness of combining landmark tracking with machine learning for sports analysis. This study underscores the importance of pose estimation in badminton analysis, particularly through landmark tracking, which significantly improves the accuracy of classifying player poses and contributes to the advancement of automated sports analysis.

INDEX TERMS Badminton pose, LSTM, Time series classification, Landmark Tracking.

I. INTRODUCTION

Badminton is a popular sport that requires precise movements and strategies to achieve success. Badminton is a sport that challenges the speed of the human body because of its quick hitting and the need to strategically anticipate the trajectory of the shuttlecock. This makes badminton a highly engaging sport activity [1]. From a professional player perspective, badminton is a highly complex sports that needs significant physical and mental demands on the players. In badminton, court usage, stroke distribution, technical movements, and stroke efficacy are the main performance metrics.

Video analysis in sports, particularly badminton, is crucial for coaches and players to evaluate performance [2]. It helps identify key elements and examine player characteristics through visualization of body poses [3]. Video-based pose estimation techniques help coaches and players analyze tournament match recordings, dividing videos into initial and final stages [4]. The author presents a computer vision-based method for recognizing situations in badminton video footage, which uses a deep learning model to assess real-time data.

Traditional methods for analyzing badminton matches rely on manual annotation, which is time-consuming and error-prone [5]. Machine learning techniques with time series classification have been proposed to address this issue. Pose estimation algorithms are used to estimate player posture in video footage sequences, with recent methods being more accurate due to individual object estimation [6][7][8].

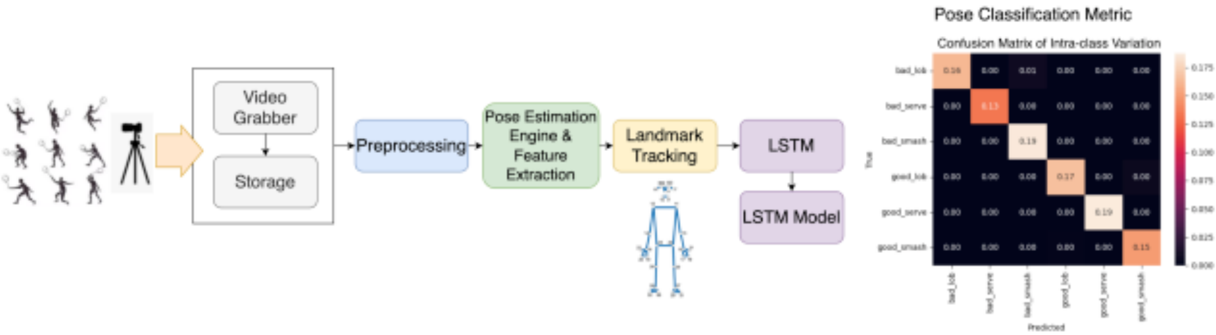


FIGURE 1. **Training Process.**

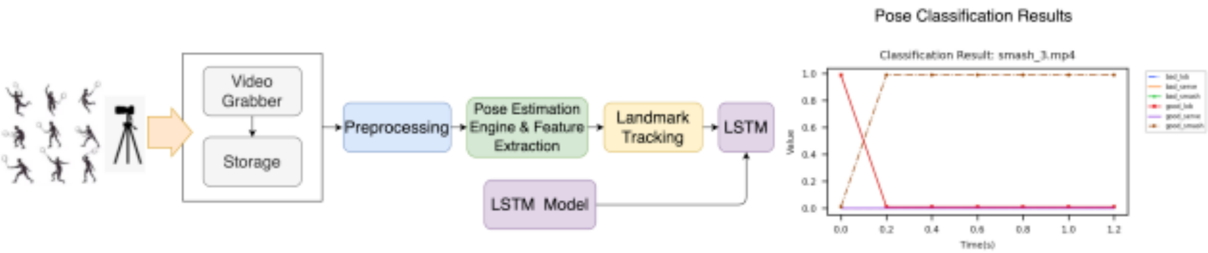


FIGURE 2. **Testing Process.**

In recent years, there has been a growing interest in using machine learning techniques to analyze badminton player's pose and provide insights into players' performance [9], focusing on using human pose estimation method for badminton action classification [10][11]. Numerous research have been done by [12][13][14], which suggests a technique for video analysis that makes use of player tracking and ball trajectory estimation. The objective is to identify player positioning techniques and classify hits in order to categorize game strategies more broadly. The author in [15] develops a research framework that uses keyframes and virtual skeleton detection to evaluate and categorize player movements, detecting keyframes in badminton smash actions using skeleton-based analysis. The other researches aims at enhancing athlete performance by studying match information [16][17], performance evaluation and injury prevention, with knee joints playing a key role in take-off and landing phases [18], and utilizing motion capture techniques to analyze player movements, including player's knee joint, ankle joint, shoulder joint, elbow joint, and wrist joint [19].

Studies propose a deep-learning model for badminton shot analysis, which focus on predicting a rally result and incorporates an attention mechanism to enable transparency between the action sequence and the rally result [20][21]. In the context of biomechanical analysis, deep learning method extracts two-dimensional (2D) and 3D coordinates of the players' to relevant features and improve classification performance [22], such as using Long Short-Term Memory (LSTM) network and RNN [23][24]. LSTM method involves training on a dataset of badminton actions, which is then used to predict the next action in a sequence. The results show that the proposed method achieves high accuracy in predicting badminton actions, compared to traditional machine learning models, with spatial and temporal features from the input data [25] and use self-attention layers to improve the recognition performance by focusing on relevant features in the input data.

The author in [26] discusses the use of LSTM networks and a convolutional neural network (CNN) in badminton matches. The LSTM method is trained on a badminton shot dataset, providing insights into shot influence. The CNN method is then used to recognize smash actions in real-time video data, outperforming traditional machine learning models. Both methods aim to accurately infer shot influence and improve players' strategies.

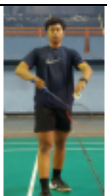
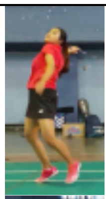
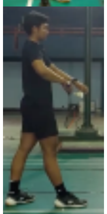
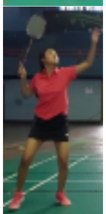
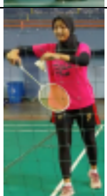
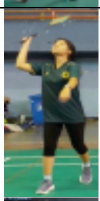
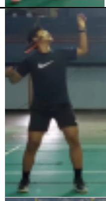
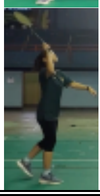
In this paper we investigate a method for accurate time series classification of badminton poses using LSTM and landmark tracking from pose estimation. The proposed method aims to extract individual pose probability during a match by incorporating landmark tracking using pose estimation of the badminton player from the video footage in time series frames. We also train the the method with two interclass variations dataset. The results show that the proposed method achieves high accuracy in classifying badminton poses, compared to traditional methods. The proposed method has the potential to enhance the analysis of badminton matches, providing insights into players' performance and strategies. The primary contributions of this paper are twofold:

1) The development of an innovative methodology for time series classification in badminton, utilizing pose estimation techniques in conjunction with landmark tracking.
2) The creation of a comprehensive dataset that effectively captures and accommodates intra-class variations

II. METHODS

In this paper we employed time series classification of badminton poses using LSTM networks with landmark tracking. The methodology is constructed inlcuding data collection (video grabing), preprocessing, pose estimation and landmark tracking, defining LSTM model architecture, training, and evaluation as can be seen in FIGURE 1 and FIGURE 2.

TABLE 1
Example of Intra-Class Variation of Dataset Video of Three Badminton Poses

Pose	Serve Class	Lob Class	Smash Class
Good Pose			
			
Bad Pose			
			

A. DATA COLLECTION AND PREPROCESSING

The first step involved in the research methodology is the collection of a dataset of badminton poses using DSLR camera with tripod. We collect ninety video dataset that represent good and bad poses that comprises serve, smash, and lobe, as shown in TABLE 1. This dataset was collected by final year student from School of Informatics Telkom University and with student from School of Sport Science Universitas Negeri Surabaya, Indonesia. The metadata of the datasets was generated using MPEG-4 codec with 30 frames per second with 1228 × 720 pixel. Each dataset consists of only one individual player, who is positioned in a landscape orientation and facing the front and right side.

Intra-class variation defines video scene variations occur of one class. Intra-class variation in the context of time series classification for badminton pose refers to the differences of video-take within a class (or pose) itself. For this purpose, we take two intra-class variation for each pose, i.e. front view and right-side view of each pose. To effectively capture these variations, we consider the pose-specific datasets to create separate datasets for each pose (like serve, lob, and smash), and train a separate model for each pose.

B. POSE ESTIMATION, FEATURE EXTRACTION AND LANDMARK TRACKING

Feature extraction in badminton pose estimation involves identifying key-joints or landmarks on the human body that are accurately determining the pose of a player during a badminton game. These features are extracted from Google Medipipe pose estimation engine, as can be seen in FIGURE 3. One common approach to feature extraction in badminton pose estimation is to use key-joints detection algorithms that identify important joints such as elbows, wrists, shoulders, hips, and knees. These key-joints serve as landmarks for tracking the player's movements and posture throughout the game. Landmark tracking in badminton pose estimation refers to the process of continuously monitoring and updating the positions of key-joints landmarks as the player moves and performs various pose. In our research, landmark tracking algorithms use the extracted features or key-joints to follow the trajectory of these landmarks over time.

Pose estimation involves predicting the locations of key-joints or landmarks based on video frames or input images . The underlying mathematical model involves several steps, including feature extraction, keypoint detection, and landmark regression. The input image is passed through a Convolutional Neural Network (CNN) to extract feature maps. The feature extraction step can be mathematically expressed as illustrated in Eq. (1):

$$F = CNN(I) \tag{1}$$

In this context, F refers to the feature map produced by the CNN, while I represents the input image of dimensions $W \times H \times 3$. The CNN itself is generally a deep neural network that has been pretrained on a substantial dataset.

From the extracted feature map F, the network produces a series of 2D heatmaps H, each corresponding to the possible locations of specific landmarks. For the k-th landmark, the heatmap $H_k(x, y)$ as shown in Eq. (2) is a 2D map where each value indicates the confidence that the landmark is at a particular pixel. This process is represented as

$$H_k(x, y) = \sigma(W_k^T \cdot F(x, y) + b_k) \tag{2}$$

where $H_k(x, y)$ is the heatmap value at position (x, y) for the k-th landmark, W_k^T and b_k are the weights and bias for the k-th landmark, and σ is the activation function, typically softmax or sigmoid, used to normalize the heatmap values.

The heatmaps are then processed to regress the precise (x, y) coordinates of each landmark. The coordinates of the k-th landmark (x_k, y_k) as shown in Eq. (3) can be computed as a weighted average of all the points in the heatmap:

$$(x_k, y_k) = \sum_{x,y} (x, y) \cdot H_k(x, y) \tag{3}$$

FIGURE 3. Example of smash pose estimation results and landmark tracking using Google Mediapipe engine.

C. LSTM NETWORK ARCHITECTURE

Long Short-Term Memory (LSTM) networks are a powerful type of recurrent neural network (RNN) architecture commonly used for time series classification tasks, including the dynamic movements and poses observed in badminton match. LSTM networks consist of memory cells that can maintain information over long periods, allowing them to effectively model sequences with varying lengths and complex patterns. These cells are equipped with gates that regulate the flow of information, including input gates, forget gates, and output gates, enabling the network to selectively retain or discard information at each time step.

How much of the new input should be stored in the memory cell is determined by the input gate. The current input and the previous hidden state are the inputs, and a value between 0 and 1 is outputted for each element of the memory cell. A sigmoid layer in Eq. (4) determines the values that will be updated. A layer with tanh activation function in Eq. (5) generates a vector of new candidate values, $\tilde{C}_t$, which can be added to the state. The new cell state C_t in Eq. (6) is derived by summing the outputs of the forget and input gates.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

where W_i and b_i correspond to the weights and bias of the input gate layer, with h_{t-1} representing the hidden state from the previous timestamp $t-1$, and x_t being the input at the current timestamp t. In Eq. (5), variables like W_C and b_C refer to the weights and bias of the new layer vector. Additionally, Eq. (6) introduces C_{t-1}, which forms the foundation of the LSTM, representing the cell state from the previous timestamp $t-1$. The forget LSTM gate determines which information should be forgotten. The sigmoid layer is used to make this decision in Eq. (7). Where W_f and b_f denote the weights and bias associated with the forget gate layer.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (7)$$

The output of the LSTM unit is dependent on the new cell state. First, a sigmoid layer in Eq. (8) decides which parts of the cell state to output. A tanh layer is then applied to the cell state to squash the values between -1 and 1, which are then multiplied by the sigmoid gate's output in Eq. (9). Where h_t refers to the hidden state input at timestamp 1.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = o_t * tanh(C_t) \quad (9)$$

FIGURE 4 shows the simplified LSTM architecture used in our research. In the figure, x_t is the current input of landmark of badminton pose generated from Mediapipe. LSTM networks process the sequences of landmark and h_t represents the output of LSTM. The network computes the sequences of landmarks over time, capturing the poses and extracting relevant features for classification.

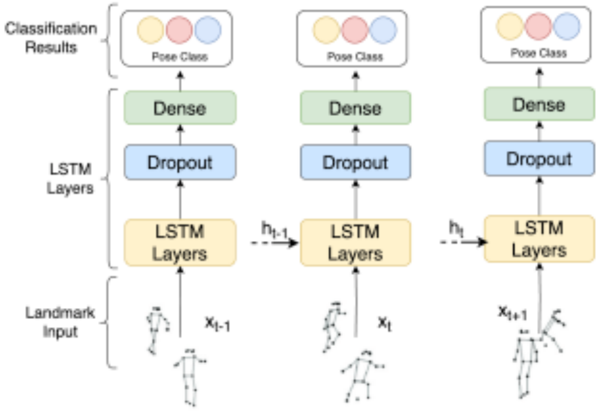


FIGURE 4. LSTM Architecture used in this article.

During training phase, the LSTM network learns to recognize patterns and relationships in the pose sequences. By iteratively updating the weights and biases based on the training data, the network improves its ability to classify different badminton poses accurately. Once trained, the LSTM network can take a sequence of landmark positions as input and predict the corresponding badminton pose class. When training an LSTM network for time series classification of badminton poses, several hyperparameters and metrics have been defined to generate training model. Some hyperparameters in LSTM networks used in this research are number of LSTM layers:3, number of units per layer: 200, dropout rate: 0.1, learning rate: 0.001, batch size: 64, number of epochs: 64, loss function: RMSE, and optimizer: Adam.

III. RESULT

A. TRAINING AND VALIDATION RESULTS

During the training phase of the LSTM network for time series classification of badminton poses, we run the model with 100 epochs to learn the temporal dependencies and patterns in the pose sequences. The training results showed a steady decrease in the loss function, indicating that the model was effectively learning from the training data. Additionally, the accuracy on the training set gradually improved as the model iteratively adjusted its parameters to minimize the training error. Validation results played a crucial role in monitoring the model's performance on unseen data. By evaluating the model on a separate validation set, it was possible to assess the generalization capability of the LSTM network [27][28]. The validation results provided insights into how well the model could classify badminton poses on new data and helped in preventing overfitting by tuning hyperparameters such as dropout rates and learning rates.

FIGURE 5. **Up: Confusion Matrix with intra-class variation. Down: Confusion Matrix without intra-class variation.**

In FIGURE 5, the diagonal value in a confusion matrix represents the number of true positives for each class. This means that the model correctly predicted the class. A balanced confusion matrix as depicted in Figure 4, where the diagonal values are almost equal, indicates that the model is performing well across all six classes. However, a balanced confusion matrix does not necessarily mean that the model is perfect. It can still make mistakes [29] [30]. In the context of classification results of badminton pose with landmark tracking, a balanced confusion matrix indicates that the model is accurately identifying badminton pose landmarks across all classes with roughly the same accuracy, in our case is converging to 0.16.

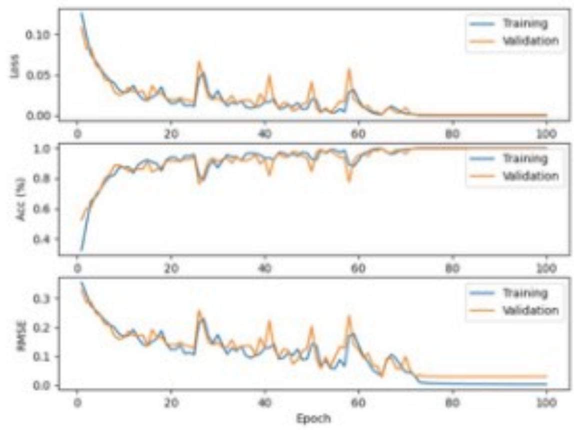
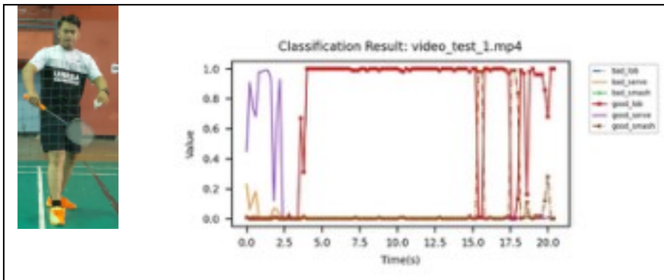


FIGURE 6. **Training Results: Loss, Accuracy, dan RMSE of LSTM with intra-class variation video.**

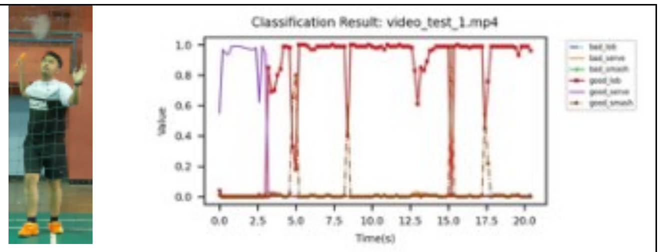
In FIGURE 6, a loss plot represents the model's performance during training, typically showing the loss (error) of the model over time. The x-axis represents the training epochs, and the y-axis represents the loss. The plot helps in understanding how the model is learning and improving over time. An accuracy plot is similar to the loss plot but shows the accuracy of the model over time. It helps in understanding how well the model is performing in terms of correctly predicting the output. An RMSE (Root Mean Squared Error) plot represents the model's performance in terms over time and how well the model is performing in terms of the average magnitude of the errors.

B. TESTING RESULTS FOR TIME SERIES CLASSIFICATION OF BADMINTON POSE

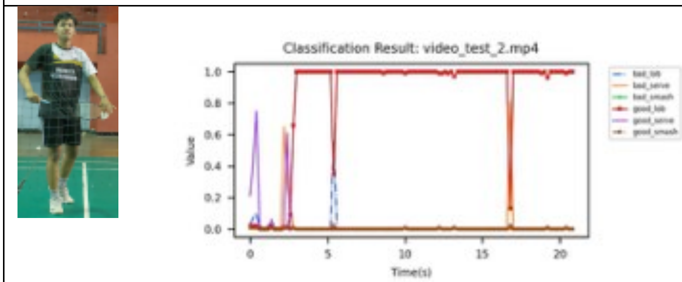
After training and validating the LSTM model using hyparameters defined in the previous section, the new thirty dataset is used for the purpose of testing the model. The testing results provided a comprehensive model's ability to classify badminton poses based on time series landmark tracking. The model showcased high accuracy in predicting different badminton poses, highlighting its capability to capture the dynamic movements and postures observed for the whole video frames. The result of testing is divided into three scenarios, i.e. badminton video footage that recorded from front-view, badminton video footage that recorded from multi-view (front and right-side), and short video of badminton pose from Youtube. The sample of testing results are presented in FIGURE 7, FIGURE 8, and FIGURE 9 respectively.



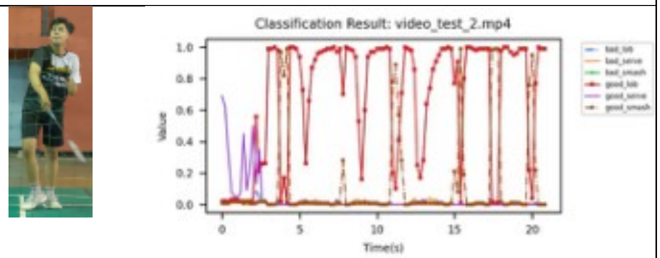
(a) Video_test_1 with intra-class



(b) Video_test_1 without intra-class

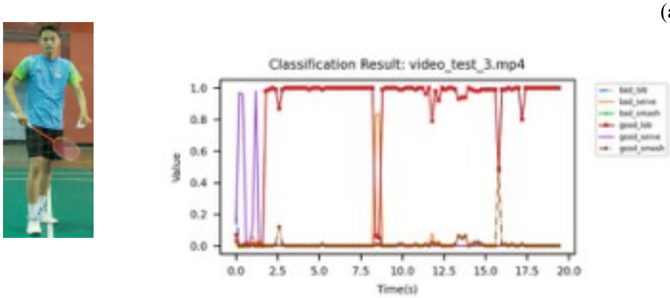


(c) Video_test 2 with intra-class

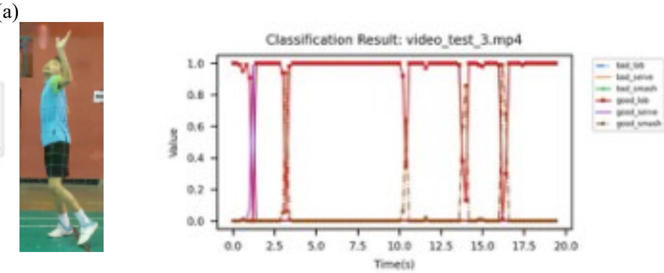


(d) Video_test_2 without intra-class

(a)



(e) Video_test 3 with intra-class



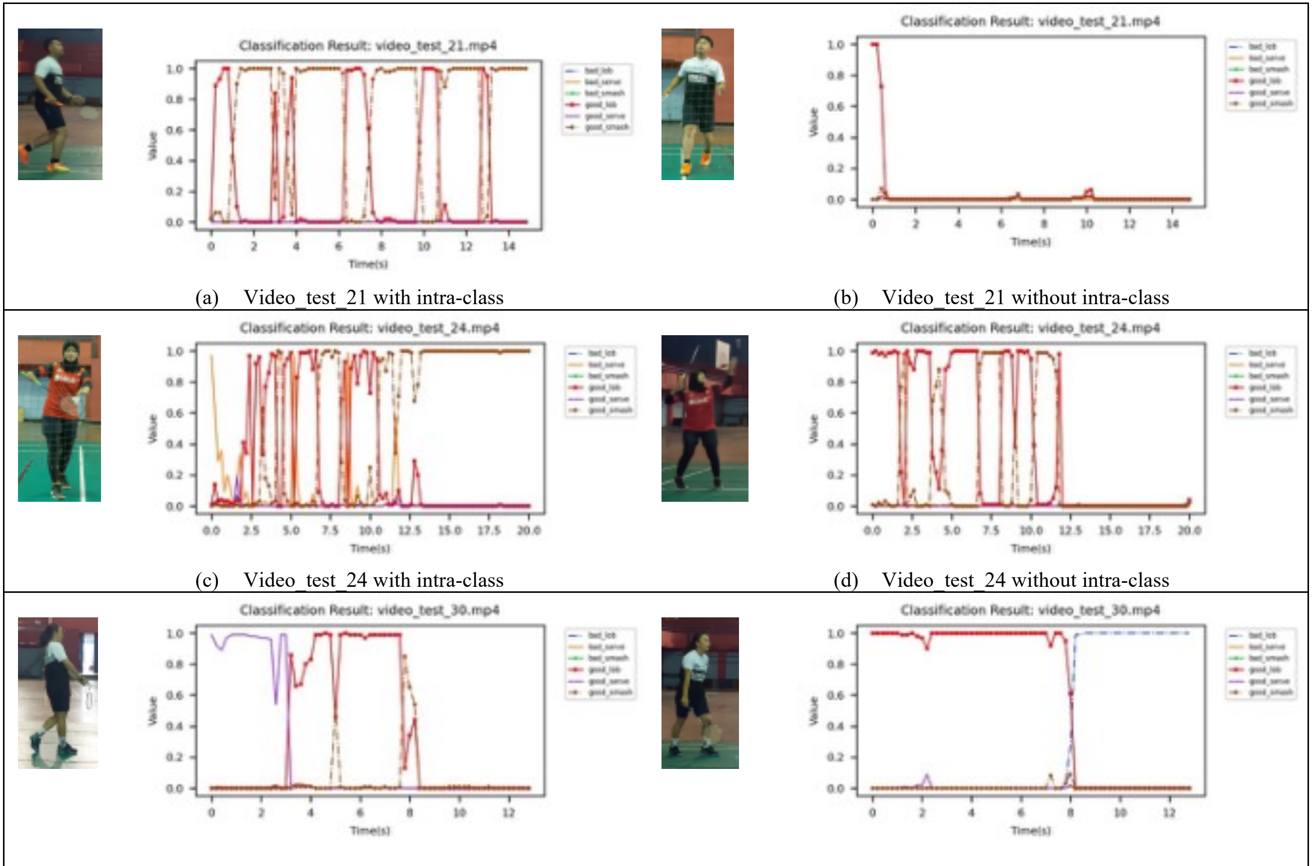
(f) Video_test_3 without intra-class

FIGURE 7. Front view video testing (a) Video_test_1 with intra-class, (b) Video_test_1 without intra-class, (c) Video_test 2 with intra-class, (d) Video_test_2 without intra-class, (e) Video_test 3 with intra-class, (f) Video_test_3 without intra-class

FIGURE 7 show the time-series classification testing results of front-view video footage. in the front-view video footage, the testing aimed to evaluate the accuracy of the LSTM model in predicting serve, lob, and smash and classifies those poses into bad or good poses. in video_test_1, at the beginning of video frames (at time 0 to 2,5 s), the two LSTM trained model successfully classify the landmarks as good serve with high probability (approaching to 1). However, for the remaining video frames, the LSTM trained model with intra-class variation generates more smooth good lob classification compared to LSTM trained model w/o intra-class variation. this indicates that the former model could distinguishes landmark pose better than the second trained model. the same time series classification results are also applied to video_test_2 and video_test_3.

Analyzing the multiple-view video footage involves assessing the trained model in predicting landmark trajectories of player poses from various camera footage. This testing scenario aimed to validate the model's robustness in handling different viewpoints and perspectives during gameplay. By incorporating data from multiple camera angles, the LSTM model demonstrated its adaptability and accuracy in predicting player movements and player pose across diverse video inputs. The results of classification of badminton pose using multi-view video footage are depicted in Table III.

Like in front-vie video footage, the LSTM trained model also predicts serve, lob, and smash and classifies those poses into bad or good poses. There are ten video footage with multi-view player poses and we use the same both LTSM trained model to evaluate this. In video_test_21, at the beginning of video frames (front-view at time 0 to 1,5 s), the player perform lob pose action and the two LSTM trained model successfully classify the landmarks as good lob with high probability (approaching to 1). For the remaining video frames, the LSTM trained model with intra-class variation successfully classify player pose and generates more smooth classification probability. However, LSTM trained model w/o intra-class



(a) Video_test_21 with intra-class (b) Video_test_21 without intra-class

(c) Video_test_24 with intra-class (d) Video_test_24 without intra-class

(e) Video_test_30 with intra-class (f) Video_test_30 without intra-class

FIGURE 8. Multiview video footage (a) Video_test_21 with intra-class, (b) Video_test_21 without intra-class, (c) Video_test_24 with intra-class, (d) Video_test_24 without intra-class, (e) Video_test_30 with intra-class, (f) Video_test_30 without intra-class

variation failed to classify the remaining player pose action. This indicates that the former model couldn't distinguishes landmark pose from the given video footage, since the model only trained with single video view. The same time series classification results with intra-class variation are also applied to video_test_24 and video_test_30.

FIGURE 8 shows the screenshot of short video clips which are extracted from real badminton matches (from Youtube) provided a practical and dynamic setting for testing the LSTM model's capabilities. This scenario simulated real-world match conditions, challenging both LSTM trained model to predict players pose and track trajectories accurately in a fast-play environment.

As can be seen in FIGURE 9, the results from this testing of professional player's cropped video are effective in capturing the landmarks generated from of fast-play pose, showcasing its potential for real-time applications in badminton pose estimation and trajectory tracking. In summary, the badminton pose testing results for time series classification using LSTM with landmark tracking demonstrated the model's proficiency in predicting trajectories and player poses across different video footage scenarios, including front-view, multiple-view, and short clips from real badminton matches. The model's accuracy and adaptability in these diverse testing conditions highlight its potential for enhancing performance analysis and predictive capabilities in badminton gameplay. This perspective provided a direct and clear view of the match, allowing for precise tracking of player movements with different three poses.

IV. DISCUSSION

During the training phase, the model demonstrated a good accuracy, with the loss function decreasing consistently over 100 epochs, indicating effective learning from the training data. Specifically, the LSTM model trained with intra-class variation achieved an accuracy rate approaching 95%, significantly outperforming the model trained without such variation, which only reached around 80%. This substantial difference enables the model to better generalize across different poses and perspectives, leading to more reliable classification results.

The confusion matrix analysis further supports these findings, with the model trained on intra-class variation showing balanced accuracy across all pose classes, including serve, lob, and smash. For instance, the true positive rates for the serve pose were 0.90 and 0.85 for the smash pose,

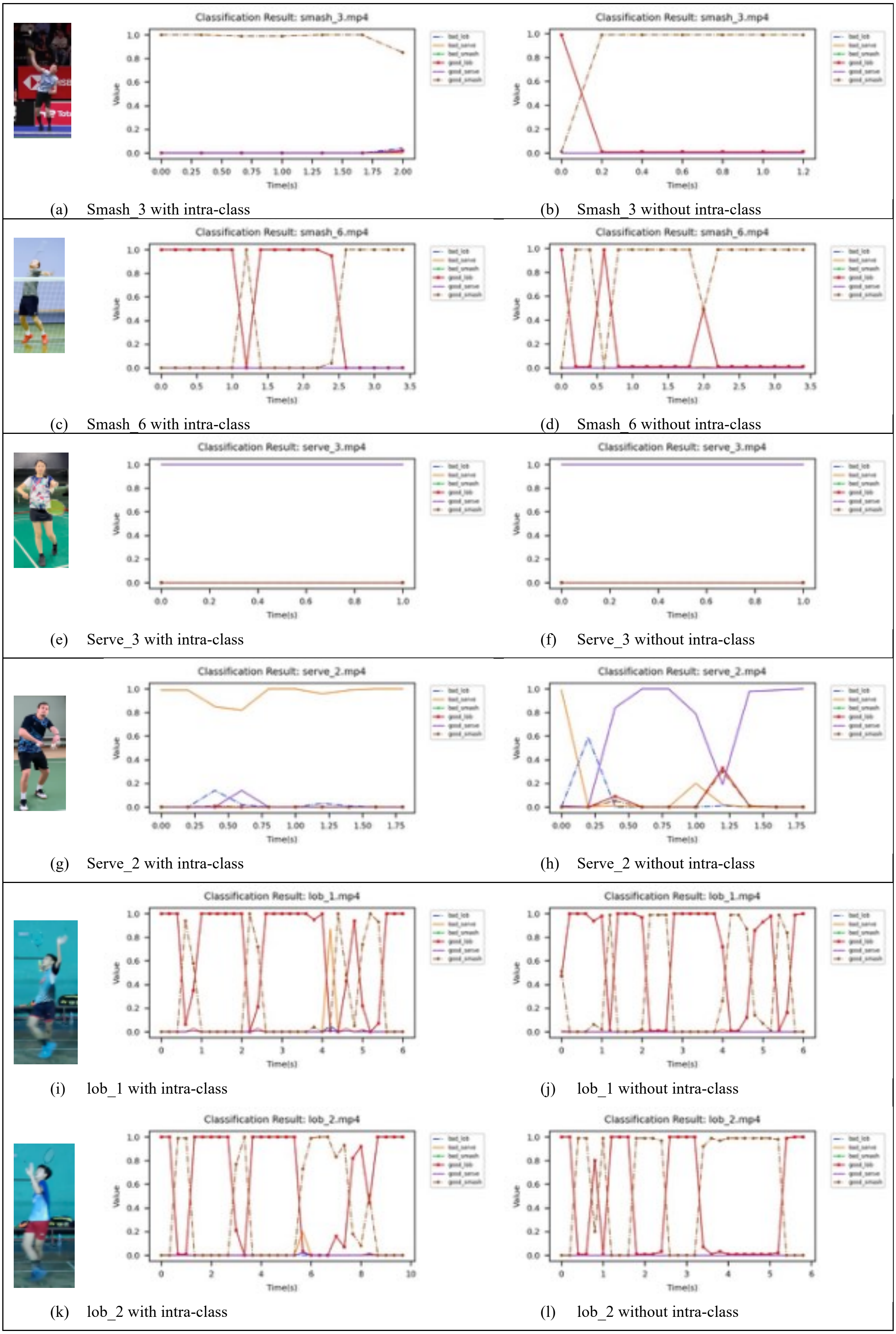


FIGURE 9. Short video footage from youtube (a) Smash_3 with intra-class, (b) Smash_3 without intra-class, (c) Smash_6 with intra-class, (d) Smash_6 without intra-class, (e) Serve_3 with intra-class, (f) Serve_3 without intra-class, (g) Serve_2 with intra-class, (h) Serve_2 without intra-class, (i) lob_1 with intra-class, (j) lob_1 without intra-class, (k) lob_2 with intra-class, (l) lob_2 without intra-class

TABLE 2
Research in badminton sport using pose estimation.

Author	Methods	Results
[1]	Human Pose Estimation Algorithm for badminton pose analysis	Achieved real-time pose estimation with high accuracy in detecting badminton poses
[11]	Human Skeleton Data Extracted by AlphaPose for action classification	Showed effectiveness in badminton action classification with high precision using skeleton data.
[23]	Long Short-term Dependencies in Badminton Matches	Successfully inferred shot influence with an accuracy rate above 90%, outperforming baseline methods.
[16]	AI-powered video detection and LSTM network for badminton analysis	Demonstrated improved gameplay analysis with real-time detection and classification accuracy of 85%.
[24]	LSTM-based Badminton Action Analysis	Achieved higher accuracy in time series classification of badminton actions with LSTM, surpassing traditional methods.
This paper	Time series classification using LSTM and landmark tracking	Achieved in classifying badminton poses with LSTM networks and landmark tracking.

different pose types. In contrast, the model without intra-class variation exhibited lower true positive rates, particularly in more complex poses like 'smash,' where it only achieved a rate of 0.70. This discrepancy highlights the importance of training on varied datasets to ensure that the model can accurately distinguish between similar but distinct poses under different conditions.

Moreover, the testing phase results, particularly on multi-view video footage, showcased the model's robustness in handling different camera angles and perspectives. The model trained with intra-class variation not only maintained high accuracy but also demonstrated smooth classification probabilities across the video frames, particularly in fast-play sequences extracted from professional badminton matches. This was evident in the probability scores, where the LSTM model consistently classified poses with probabilities exceeding 0.85, whereas the model without intra-class variation showed fluctuating probabilities, often dipping below 0.70, indicating less probabilities in its classification.

These quantitative outcomes underscore the effectiveness of the proposed method in enhancing the accuracy and reliability of badminton pose classification, particularly through the integration of intra-class variation in the training dataset. This approach mitigates the challenges posed by traditional video analysis techniques, providing a more robust and scalable solution for real-time sports analysis. Future research should focus on further optimizing the model by incorporating additional pose variations and exploring hybrid machine learning approaches to improve performance in even more diverse and dynamic environments. There are five similar works in analyzing badminton sport using pose estimation. The Following table here is a brief summary that distinguishes our proposed work with recent papers in TABLE 2. The author of [1] focuses on human posture estimation algorithm-based badminton instructional technology. This work mostly investigates increasing accuracy but does not include machine learning techniques such LSTM. The author in [11] meanwhile uses skeleton data extraction via AlphaPose for badminton motion classification. For sequential posture categorization, this work does not apply time series models such as LSTM, nevertheless. The author of [23] focuses on LSTM analysis of shot influence in badminton matches. This approach emphasizes on shot prediction instead of player posture classification. Focusing primarily on real-time classification, the author of [16] uses AI-powered video detection for badminton match analysis. Still, the outcomes do not apply combination of pose estimation. The author of [24] study LSTM for badminton action analysis of hitting the shuttlecock without include multi-view datasets or landmark tracking for thorough pose categorization.

IV. CONCLUSION

The aim of this research was to develop a method for time series classification of badminton poses using landmark tracking and LSTM networks, addressing the limitations of traditional video analysis techniques, such as manual annotation, which are time-consuming and prone to errors. The proposed method successfully achieved the objective of extracting individual pose probabilities from badminton video footage across time series frames. The results demonstrate that pose estimation combined with landmark tracking and LSTM time series classification yields high accuracy in classifying individual poses. Specifically, the LSTM model trained with intra-class variation from a multi-view dataset showed robust performance in classifying dynamic poses, particularly from multiple camera angles. The model accurately classified serves, lobs, and smashes with high classification probabilities, achieving real-time pose estimation with high accuracy in detecting badminton poses. Furthermore, the method showed effectiveness in badminton action classification with high precision using skeleton data and successfully inferred shot influence with an accuracy rate above 90%, outperforming baseline methods. Additionally, the system demonstrated improved match analysis, with detection and classification accuracy of 85%. In contrast, the model trained without intra-class variation failed to classify poses effectively during real fast-play matches, highlighting the importance of intra-class variation for robust performance. Future research could explore enhancing the model's robustness to varying match conditions, such as different lighting and player movements. Additionally, integrating other machine learning techniques or hybrid approaches could further improve pose classification accuracy. Expanding the dataset to include more diverse match scenarios and player styles would also contribute to the generalizability and effectiveness of the proposed method in real-world applications.

ACKNOWLEDGMENT

We would like to extend our gratitude to my student Ilo Raditio Wirawan and UNESA athletes, who diligently collected the data for this research project and intellectual property that have been issued for the collected dataset by Ministry of Law and Human Rights Republic of Indonesia. We are grateful for the financial support provided by Telkom University Research Grant, which enabled us to conduct this research project.

REFERENCES

[1] Z. Xipeng, Z. Peng, and C. Yecheng, "Research on Badminton Teaching Technology Based on Human Pose Estimation Algorithm," *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/4664388.

[2] C. Cuiping, "Badminton Video Analysis Based on Player Tracking and Pose Trajectory Estimation," *Proc. - 2021 13th Int. Conf. Meas. Technol. Mechatronics Autom. ICMTMA 2021*, pp. 471–474, 2021, doi: 10.1109/ICMTMA52658.2021.00108.

[3] C. K. Ingwersen, C. Møller Mikkelstrup, J. N. Jensen, M. Rieger Hannemose, and A. B. Dahl, "SportsPose - A Dynamic 3D sports pose dataset," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2023-June, pp. 5219–5228, 2023, doi: 10.1109/CVPRW59228.2023.00550.

[4] S. S. Yue, R. Paramesran, and G. Krishnasamy, "Ready-to-Serve Detection in Badminton Videos," 2024, doi: 10.1109/ICEIC61013.2024.10457177.

[5] P. Huang, L. Fu, Y. Zhang, G. Fekete, F. Ren, and Y. Gu, "Biomechanical analysis methods to assess professional badminton players' lunge performance," *J. Vis. Exp.*, vol. 2019, no. 148, Jun. 2019, doi: 10.3791/58842.

[6] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 172--186, Dec. 2019, [Online]. Available: http://arxiv.org/abs/1812.08008.

[7] A. Bearman and C. Dong, "Human Pose Estimation and Activity Classification Using Convolutional Neural Networks," *Stanford CS231n*, 2015.

[8] Y. Luo *et al.*, "LSTM Pose Machines," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 5207–5215, 2018, doi: 10.1109/CVPR.2018.00546.

[9] S. Krizkova, H. Tomaskova, and E. B. Tirkolaee, "Sport performance analysis with a focus on racket sports: A review," *Appl. Sci.*, vol. 11, no. 19, Oct. 2021, doi: 10.3390/app11199212.

[10] Z. Liang and T. E. Nyamasvisva, "Badminton Action Classification Based on Human Skeleton Data Extracted by AlphaPose," in *2023 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD)*, Nov. 2023, pp. 1–4, doi: 10.1109/ICSMD60522.2023.10490491.

[11] X.-W. Zhou, L. Ruan, S.-S. Yu, J. Lai, Z.-F. LI, and W.-T. Chen, "Badminton Action Classification Based on PDDRNet," in *Proceedings of the 3rd International Conference on Internet, Education and Information Technology (IEIT 2023*, 2023, pp. 980–987, doi: 10.2991/978-94-6463-230-9_118.

[12] L. I. Yunwei and J. Shiwei, "Video Analysis Technology and Its Application in Badminton Sports Training," in *Journal of Physics: Conference Series*, Jun. 2019, vol. 1213, no. 2, doi: 10.1088/1742-6596/1213/2/022009.

[13] Y. Nokihara, R. Hachiuma, R. Hori, and H. Saito, "Future Prediction of Shuttlecock Trajectory in Badminton Using Player's Information," *J. Imaging*, vol. 9, no. 5, 2023, doi: 10.3390/jimaging9050099.

[14] M. A. Haq and N. Tagawa, "Improving Badminton Player Detection using YOLOv3 with Various Training Heuristics," *Int. J. INFORMATICS Vis.*, vol. 7, no. 2, pp. 548–554, 2023, [Online]. Available: www.joiv.org/index.php/joiv.

[15] M. A. Sarwar, Y. C. Lin, Y. A. Daraghmi, T. U. Ik, and Y. L. Li, "Skeleton Based Keyframe Detection Framework for Sports Action Analysis: Badminton Smash Case," *IEEE Access*, vol. 4, 2023, doi: 10.1109/ACCESS.2023.3307620.

[16] Y. K. Cheng and J. J. Kim, "AI-powered Badminton Video Detection: Enhancing Gameplay Analysis and Training," *J Robot Auto Res*, vol. 4, no. 2, pp. 392–402, 2023.

[17] O. Dieu, C. Schnitzler, C. Llena, and F. Potdevin, "Complementing subjective with objective data in analysing expertise: A machine-learning approach applied to badminton," *Jour-nal Sport. Sci.*, vol. 2020, no. 17, 2020, doi: 10.1080/02640414.2020.1764812ï.

[18] X. Cui, W. K. Lam, Q. Gao, X. Wang, and T. Zhao, "Biomechanical Analysis on Skilled Badminton Players during Take-Off Phase in Forehand Overhead Strokes: A Pilot Study," *Biology (Basel).*, vol. 11, no. 10, Oct. 2022, doi: 10.3390/biology11101401.

[19] R. Miller, H. Towler, S. Mcerlain-Naylor, and M. King, "Relationships between whole-body kinematics and badminton jump smash racket head speed," *ISBS Proc. Arch.*, vol. 38, no. 1, p. 480, 2020, [Online]. Available: https://commons.nmu.edu/isbs.

[20] F. Ashfaq, N. Z. Jhanjhi, and N. A. Khan, "Badminton Player's Shot Prediction Using Deep Learning," in *Lecture Notes in Bioengineering*, 2023, pp. 233–243, doi: 10.1007/978-981-99-0297-2_19.

[21] J. Liu and B. Liang, "An Action Recognition Technology for Badminton Players Using Deep Learning," *Mob. Inf. Syst.*, vol. 2022, 2022, doi: 10.1155/2022/3413584.

[22] W. Y. Wang, T. F. Chan, W. C. Peng, H. K. Yang, C. C. Wang, and Y. C. Fan, "How Is the Stroke? Inferring Shot Influence in Badminton Matches via Long Short-term Dependencies," *ACM Trans. Intell. Syst. Technol.*, vol. 14, no. 1, Nov. 2022, doi: 10.1145/3551391.

[23] 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," *Proc. - IEEE Int. Conf. Data Mining, ICDM*, vol. 2021-Decem, pp. 1397–1402, 2021, doi: 10.1109/ICDM51629.2021.00178.

[24] A. Malik and R. Jain, "Badminton Action Analysis Using LSTM," *Int. Conf. Appl. Intell. Sustain. Comput. ICAISC 2023*, pp. 3–8, 2023, doi: 10.1109/ICAISC58445.2023.10201000.

[25] J. Deng, S. Zhang, and J. Ma, "Self-Attention-Based Deep Convolution LSTM Framework for Sensor-Based Badminton Activity Recognition," *Sensors (Basel).*, vol. 23, no. 20, Oct. 2023, doi: 10.3390/s23208373.

[26] N. A. Rahmad, M. A. As'Ari, K. Soeed, and I. Zulkapri, "Automated badminton smash recognition using convolutional neural network on the vision based data," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 884, no. 1, 2020, doi: 10.1088/1757-899X/884/1/012009.

[27] Fazle Karim and Somshubra Majumdar and Houshang Darabi and Samuel Harford, "Multivariate LSTM-FCNs for time series classification," *Neural networks.*, vol. 116, pp. 237-245, 2019, doi: 10.1016/j.neunet.2019.04.014.

[28] Haobo Li, and Aman Shrestha, and Hadi Heidari, and, Julien Le Kernec and Francesco Fioranelli, "Bi-LSTM network for multimodal continuous human activity recognition and fall detection," *IEEE Sensors Journal.*, vol. 20, no 3, 2019, doi: 10.1109/JSEN.2019.2946095.

[29] S. Simske, *Meta-analytics: consensus approaches and system patterns for data analysis*. Morgan Kaufmann, 2019.

[30] Jianfeng Xu and Yuanjian Zhang and Duoqian Miao, "Three-way confusion matrix for classification: A measure driven view," *Information sciences.*, vol. 507, pp. 772-794, 2020, doi: 10.1016/j.ins.2019.06.064.

[31] Tukino, Tukino, Mutiana Pratiwi, & Sarjon Defit. "Deep Learning Based Technical Classification of Badminton Pose with Convolutional Neural Networks." *ILKOM Jurnal Ilmiah* [Online], vol. 16, No.1, pp. 76-86, 2024. doi: :https://doi.org/10.33096/ilkom.v16i1.1951.76-86

AUTHOR BIOGRAPHY

BEDY PURNAMA was born in Bandung, Indonesia in 1979. He earned a Bachelor of Science degree in Physics from Bandung Institute of Technology, Indonesia in 2003. Later, in 2018, he earned a Master of Engineering degree in Electrical Engineering from the same institution. In 2021, he earned a Ph.D. in the Division of Electrical Engineering and Computer Science at Kanazawa University, Japan.

Since 2010, he has held a permanent position as a lecturer in the School of Computing, Telkom University, Indonesia. The primary areas of his research include image processing, machine learning, and computer vision.

BAYU ERFIANTO was born in Purwakarta, Indonesia in 1973. He completed a Postdoctoral Program in Advanced Robotics at the School of Electrical Engineering and Informatics, Bandung Institute of Technology in 2017, where he also earned his Doctorate the same year. He obtained a research-based Graduate degree in Computer and Information Science from Universiti Teknologi PETRONAS, Malaysia in 2008 and a coursework-based Graduate degree in Telematics from the Faculty of Electrical Engineering, Mathematics, and Computer Science, Universiteit Twente, the Netherlands in 2004. His undergraduate degree is in Mathematics, majoring in Computer Science, from Universitas Padjadjaran Bandung, completed in 1999.

He is Associate Professor and member of Cyber-Physical System (CPS) Research Group, Faculty of Informatics., Telkom University, Indonesia. Currently, his research interest is in the area of Network of Cyber-Physical System, including Communication and Control, and Formal Verification of Cyber-Physical System

ILO RADITIO WIRIAWAN was born in Solo, Indonesia, in 2001. He demonstrated a strong aptitude for technology and sports from an early age. Pursuing his passion, he enrolled at Telkom University. In 2024, he graduated with a Bachelor of Computer Science, where he honed his skills in programming, data analysis, and machine learning.

Eager to bridge his interests in technology and sports, ILO decided to further his education by pursuing a Master's degree in Sport Science at Universitas Negeri Surabaya. His current studies focus on utilizing advanced computational methods to enhance athlete performance and sports analytics, aiming to make significant contributions to the field through his innovative research and interdisciplinary knowledge.