ILKOM Jurnal Ilmiah Vol. 16, No. 1, April 2024, pp. 76-86
Accredited 2nd by RISTEKBRIN No. 200/M/KPT/2020; E-ISSN 2548-7779 | P-ISSN 2087-1716

Research Article

Deep Learning Based Technical Classification of Badminton Pose with Convolutional Neural Networks

Tukino [a,1,*]; Mutiana Pratiwi [b,2]; Sarjon Defit [b,3]
[a] *Information System, Universitas Putera Batam, Jl. Soeprapto, Batam, 29434, Indonesia*
[b] *Information System, Universitas Putra Indonesia YPTK Padang, Jl. Raya Lubuk Begalung, Kota Padang, Sumatera Barat 25145*
[1] *tukino@puterabatam.ac.id;* [2] *mutiana_pratiwi@upiyptk.ac.id;* [3] *sarjon_defit@upiyptk.ac.id*
* *Corresponding author*

Article history: Received November 20, 2023; Revised February 23, 2024; Accepted April 02, 2024; Available online April 26, 2024

Abstract

This research aims to identify and categorize badminton strategies using a Convolutional Neural Network (CNN) model combined with BlazePose architecture and Mediapipe Pose Solution tools, yielding understandable and practical results. The challenge of finding the best mobility strategy for badminton serves as the primary motivation for this study. The research employs an image recognition and supervised learning approach to classify poses in badminton training videos. The training data comprises various photos and images representing different badminton techniques, such as Service Technique and Smash Technique. After data processing, the CNN model is trained using the training data to identify and classify poses in badminton training videos. Testing is conducted using test data, and classification accuracy is evaluated using the CNN method. The results show that the CNN model implemented alongside BlazePose and Mediapipe Pose Solution achieves significant classification accuracy, ranging from 80% to 90%. Thus, this research presents an effective and practical method for classifying badminton strategies based on poses in training videos.

Keywords: Badminton; Classification; Convolutional Neural Networks; Deep Learning; Estimation Pose

Introduction

Health and human existence are intrinsically interwoven. Without a healthy body, humans would suffer disruptions and deteriorate in physical fitness. Health, a state of balance, is determined by factors in a person's genetic make-up, environment, and way of life, including what they eat, how they exercise, how they interact with others, how they work, how they relax, and how they manage their emotions [1]. Nowadays, a lot of individuals live unhealthy lives, particularly when it comes to their eating, exercising, and sleeping patterns. Keeping a diet is crucial for managing how much food the body requires. Along with maintaining a healthy diet and engaging in other crucial activities, we also need to practice sports [2].

Badminton is a well-known sport across the world. With both national and international events, badminton has enjoyed considerable success in Indonesia. In the numbers, men's and women's singles, men's and women's doubles, and other matches are commonly played. The accomplishments attained in the contests in which badminton competitors have participated are proof that the sport has evolved incredibly swiftly in Indonesia [3]. Achievement with the aid of science and technology must be established via a planned, progressive, and ongoing process of training and development [4].

Automatic image detection disputes are still a key topic in computer vision research. The researchers developed a number of parallel methodologies and mathematical techniques to address the discrepancy between 3D object detection and 2D object detection. Computer vision is the ability to view an object in order to display digital elements and graphically gather data on a computer. Machines can perceive and handle objects in the same manner that people can thanks to computer vision [5].

One of the many methods that may be used to categorize and identify photos is the Convolutional Neural Network (CNN) approach. Convolutional neural networks are built on the MLP, a feed forward (non-repeating) neural network type (CNN) [6].

Numerous studies with good accuracy have used the CNN approach with great success. Rasywir et alstudy, .'s "Analysis and Implementation of Oil Palm Disease Diagnostics using the CNN Method," obtained an accuracy of 87% with a total dataset of 2490 oil palm pictures classified with 11 sickness categories. A research by [7] titled "Face Recognition for Bank Employee Access Using Deep Learning Using the CNN Method" using a total dataset of 5 bank employee faces and 70 facial data points for each person produced a 95% accuracy rate.

The CNN model's learning capabilities, which makes it straightforward to learn new characteristics using input data in the form of images rather than having to extract them for classification scenarios, is an advantage for spotting badminton. Based on the aforementioned characteristics, the researcher plans to use the CNN technique, which can categorize pictures using significant output and is simple to extract features from images. This is thought of as a means to categorize the different badminton techniques based on postures. Therefore, in this study, we will use the CNN model, the BlazePose architecture, and the Mediapipe Posture Solution tools to perform pose estimation in order to extract landmarks and keypoints for each input picture [8].

Computer vision-based systems have demonstrated that they are progressing, including Optical Character Recognition (OCR), medical image recognition, vehicle security, fingerprint recognition, biometrics, and others. The use of the convolutional neural network approach in deep learning is seen to be highly important when accuracy and efficiency are taken into account [9]. Below is a list of the research that make use of CNN. A system that can recognize illness and deliver information in the form of treatment options for illnesses that affect tomato leaves was developed using CNN. This was accomplished via supervised classification with 200 samples of tomato leaf photos, 160 of which served as training data and 40 as test data. According to the test results, the CNN method has average values for accuracy, precision, recall, and error rate of 97.5%, 95.45%, and 5%, respectively. These values include 95% accuracy, 90.83% precision, 90% recall, and 10% error on average when compared to SVM. The test results show that CNN performs better as a classifier in this study than SVM [10]. CNN can identify faces that are still present in digital images that are extracted frame output images from CCTV video with an accuracy level of 80% for objects that have already been registered in the database and with an accuracy level of 40% for objects that have not [11]. This is done in order to recognize CCTV video objects and produce images using the CNN method.

Artificial intelligence (AI), the capacity of a computer to simulate human intellect, enables it to carry out tasks that would otherwise be carried out by a person [12]. Machine learning, sometimes known as "machine learning," is the most popular approach since it is routinely used to imitate or replace human behavior when resolving problems [13]. The two applications of machine learning are categorization and prediction. Classification is a Machine Learning (ML) approach that enables the computer to classify or categorize items based on particular criteria, much as how people attempt to distinguish one thing from another. Machines use the ML approach for prediction in the interim to forecast results based on input data based on data that has been examined [14].

CNN, which also include a number of hidden layers and are arranged as an architecture, are a type of artificial neural networks with weights. The CNN model has several layers, including convolution, activation function, pooling layer, flatten layer, and fully connected layer. Flow of CNN is depicted in **Figure 1** [15]. Technically, each input image will be passed through a series of convolution layers with filters (Kernals), Pooling, fully connected layers (FC), and apply Softmax function to classify an object with probabilistic values between 0 and 1. This is done in order to train and test deep learning CNN models. The flowchart for how CNN processes an input picture and categorizes the objects based on values is shown below [16].

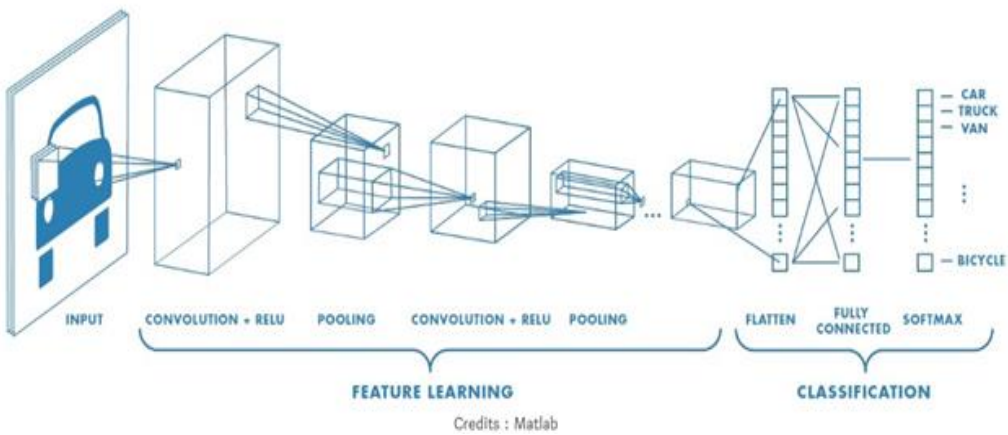

Figure 1. CNN Flow ([16])

The first layer to extract features from an input picture is convolution. Convolution learns visual characteristics from tiny input data squares, preserving the link between pixels. Two inputs, such as an image matrix and a filter or kernel, are required for this mathematical procedure [17].

One of the challenging areas of research in computer vision is human pose estimation, which tries to extract the position or spatial location of significant human body parts and joints from a given picture or video. The technique of identifying postures in a picture, known as "human pose estimation," can be carried out in 2D or 3D [18].

The link between explanatory factors (independent variables) and response variables (dependent variables), which are qualitative in nature and include two or more categories, is examined using a statistical approach called logistic regression [19].

CNN Flow is a term that could refer to several different things, depending on the context. It could mean a CNN architecture designed specifically for flow data, such as optical flow or fluid dynamics. Alternatively, it might refer to the flow of data through a CNN model during the training or inference process. [20].

CNN Flow could imply the application of CNNs to analyze or process data related to flow, whether it's in the context of computer vision, fluid dynamics, or another field where flow data is relevant. If you have a specific context or usage in mind, please provide more details for a more precise explanation. [21].

Method

A. *Population and Sample*

The population of the data consists of badminton techniques that were downloaded from the YouTube website. The three badminton strokes that were selected as examples in this study are the serve, forehand, and smash [22].

B. *Types and Sources of Data*

Data from the YouTube website that have been transformed to suit the model were the sort of data utilized in this investigation.

C. *Data Analysis Method*

The analytical approach utilized to locate postures and recognize landmarks and keypoints is the CNN using the Mediapipe pose solution method. The classification method makes use of a number of supervised learning approaches, such as logistic regression, random forest, and k-nearest neighbor [23].

D. *Research Stages*

The stages passed in this study are described in the following diagram:

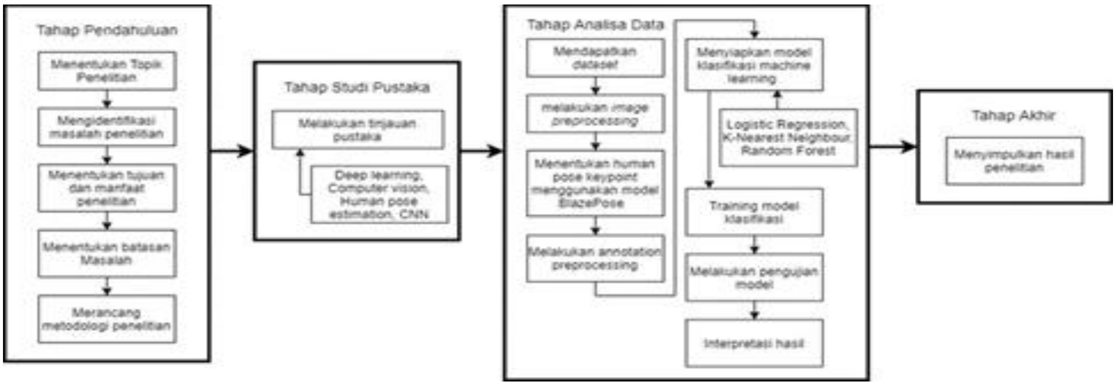

Figure 2. Research Stages

1) *Criteria for Video Selection*

In selecting videos to form the dataset, we considered several criteria to ensure a good representation of various badminton techniques. Firstly, we chose training videos that encompassed a range of fundamental techniques, such as Serving Technique and Smash Technique. We also selected videos that exhibited variations in intensity and speed of movement, as well as diverse angles and lighting conditions. Additionally, visual quality was a crucial consideration, ensuring that the videos had adequate resolution and quality to facilitate precise identification of poses.

2) *Preprocessing Steps*

After selecting the videos, preprocessing steps were conducted to prepare the dataset for model training. These steps included image normalization to ensure consistency in scale and intensity, image cropping to adjust pose sizes, and data augmentation to increase variability in the dataset. Data augmentation techniques included image rotation, horizontal and vertical shifts, as well as image flipping to create more variations in poses.

3) Data Set Division

The processed dataset was then divided into three sets: training set, validation set, and test set. This division was done with appropriate proportions to ensure accurate evaluation of the model. Typically, around 70-80% of the dataset was used for training, 10-15% for validation, and the remaining portion for testing. Each dataset was randomly selected to ensure fair representation of the entire dataset.

4) Selection of Learning Models

The selection of the CNN model for classification was based on the need to evaluate its performance in classifying badminton poses. CNN was chosen due to its proven effectiveness in image recognition tasks and its ability to capture spatial dependencies in data. Additionally, CNNs have been widely utilized in various domains for their ability to automatically learn hierarchical features from raw data. By focusing solely on the CNN model, we aimed to assess its suitability and performance specifically in the context of badminton pose recognition. CNNs have shown promising results in similar tasks, making them a suitable candidate for this study. Additionally, by exclusively evaluating the CNN model, we could gain a deeper understanding of its strengths and weaknesses in handling the complexities of badminton pose classification.

E. Dataset Design

Videos make up the bulk of the dataset that was used in this study. A badminton player's whole two-second set of motions are captured, serialized into photographs, and then combined into various batch sizes to provide the data needed to create movies by breaking them down into frame-by-frame images. The footage of each badminton player has been preprocessed to cut down on the number of frames.

Table 1. Number of Images in the Dataset

No	Videos	Number of Images
1	Service	369
2	Forehand	374
3	Smash	420
Amount		1163

Results and Discussion

A. Results

1) Dataset Collection and Data Preprocessing

Video is the first sort of data to be found. Forehand methods, service techniques, and smash techniques make up the three categories of the data. The YouTube website is used to download it. In order to make it simpler to annotate data, downloaded videos are played by a single player.

2) Data Annotation

In this work, estimating postures from each type of dataset is the process of data annotation. A CSV file containing the file name, data class, and x, y, and z coordinates of each Landmark will be the process' output. The Mediapipe Pose Solution library and the Google Colab environment are also used by the researcher throughout this procedure. Obtaining, analyzing, and performing a pose estimation on data is the initial stage in the boostrapping process.

Obtaining, analyzing, and performing a pose estimation on data is the initial stage in the boostrapping process. It is essential to provide the input image location, output image location, and output location for the csv file prior to performing the technique.

```
Bootstrapping  forehand
100%|██████████| 374/374 [01:05<00:00,  5.71it/s]
Bootstrapping  servis
100%|██████████| 369/369 [00:59<00:00,  6.19it/s]
Bootstrapping  smash
100%|██████████| 420/420 [01:12<00:00,  5.78it/s]
```

Figure 3. The Boostrapping Process

```
Number of images per pose class:
  forehand: 374
  servis: 369
  smash: 420
```

Figure 4. Number of datasets for each class

The results of the preceding boostrapping are presented in the form of class-specific CSV files and landmarked images. The classification model can be trained using these data, but they still contain outliers, needing a separate step to get rid of them. The results of the bootstrapping are shown in the following graphic.

Figure 5. Forehand Technique bootstrapping results

Table 2. Forehand technique bootstrapping csv results

No	File Name	Nose			Right Eye Inner			…	Left Foot Index		
		x1	*y2*	*z3*	*x2*	*y2*	*z2*	…	*x33*	*y33*	*z33*
1	Ser (1).jpg	609.89	286.69	-228.47	609.25	275.82	-212.31	…	606.44	591.82	334.25
2	Ser (2).jpg	638.78	201.99	-336.62	644.33	190.50	-316.60	…	713.65	738.57	28.34
3	Ser (3).jpg	669.23	199.36	-297.68	676.26	188.92	-275.56	…	752.22	664.73	294.91
4	Ser (4).jpg	641.07	226.58	-355.23	646.62	212.41	-341.27	…	765.18	661.19	397.74
5	Ser (5).jpg	634.35	239.37	-303.23	639.02	228.01	-292.13	…	802.23	639.55	152.48
6	Ser (6).jpg	655.59	255.95	-345.36	656.29	244.28	-329.39	…	811.16	639.95	304.67
7	Ser (7).jpg	714.37	345.18	-276.43	726.59	321.90	266.14	…	1071.17	972.94	185.56
8	Ser (8).jpg	759.87	344.69	-294.48	768.42	322.01	-292.47	…	1076.39	1088.72	62.35
9	Ser (0).jpg	664.44	244.82	-269.01	670.13	232.59	-258.01	…	822.72	635.26	241.27
10	Ser (10).jpg	675.24	207.29	-379.27	677.44	195.45	-354.73	…	798.61	649.08	245.02

Table 3. Service Engineering bootstrapping csv results

No	File Name	Nose			Right Eye Inner			…	Left Foot Index		
		x1	*y2*	*z3*	*x2*	*y2*	*z2*	…	*x33*	*y33*	*z33*
1	Ser (1).jpg	682.10	97.88	-327.45	692.34	83.24	-295.04	…	820.38	869.03	189.99
2	Ser (2).jpg	729.19	95.78	-353.36	747.88	99.41	-335.48	…	505.53	656.76	374.28
3	Ser (3).jpg	664.14	157.75	-286.38	678.72	160.07	-269.19	…	433.97	716.30	145.05
4	Ser (4).jpg	712.87	244.65	-413.57	720.38	245.00	-402.53	…	394.36	395.20	589.13
5	Ser (5).jpg	721.47	214.06	-305.56	729.13	219.39	-338.40	…	353.92	65.03	140.49
6	Ser (6).jpg	697.25	142.02	14.81	710.29	142.10	35.53	…	593.17	483.42	441.52
7	Ser (7).jpg	564.62	77.94	-663.80	561.89	71.15	-699.70	…	422.93	695.82	249.86
8	Ser (8).jpg	629.12	-11.14	-414.09	637.90	-12.50	-399.98	…	559.99	619.91	237.27
9	Ser (0).jpg	521.85	-36.37	-213.65	525.17	-51.68	-207.56	…	494.30	621.31	293.44
10	Ser (10).jpg	514.50	79.87	-360.37	524.79	64.41	-376.73	…	480.74	618.36	282.43

Figure 6. The results of the Smash Technique bootstrapping

Figure 7. Service Engineering bootstrapping csv results

Table 4. Results of the Smash Technique bootstrapping csv

No	File Name	Nose			Right Eye Inner			…	Left Foot Index		
		x1	*y2*	*z3*	*x2*	*y2*	*z2*	…	*x33*	*y33*	*z33*
1	Ser (1).jpg	642.59	228.73	-264.22	649.30	216.77	-254.88	…	575.22	598.40	430.25
2	Ser (2).jpg	602.74	197.07	182.71	610.28	187.77	-165.72	…	583.25	602.57	86.15
3	Ser (3).jpg	598.73	196.67	177.73	605.27	189.02	-160.52	…	589.58	601.12	211.32
4	Ser (4).jpg	593.29	196.14	196.94	599.27	189.29	-179.45	…	595.83	588.46	193.96
5	Ser (5).jpg	605.96	195.44	184.97	612.22	189.63	-165.34	…	599.64	597.26	157.80
6	Ser (6).jpg	442.57	216.33	-277.80	445.60	207.43	-259.17	…	454.16	560.15	391.62
7	Ser (7).jpg	683.15	518.62	32.58	686.27	521.50	-28.10	…	640.59	448.20	-127.05
8	Ser (8).jpg	762.24	375.22	10.21	755.68	367.22	0.52	…	870.41	689.21	-176.14
9	Ser (0).jpg	967.77	353.61	25.39	963.38	344.48	-15.06	…	1028.75	693.99	131.72
10	Ser (10).jpg	886.84	353.12	2.88	881.59	345.05	15.12	…	893.11	705.86	-222.54

3) Annotation Data Preprocessing

There are several reasons why this study's findings are considered outliers, including:

a. If there are any prediction mistakes, the outliers will be taken out of the dataset.
b. Incorrect initial classification: After the data annotation process, each sample is classed against the database; if the samples do not fall into the same category, the samples are labeled as outliers.

```
print('Number of outliers: ', len(outliers))

Number of outliers:  146
```

Figure 8. Number of Outliers

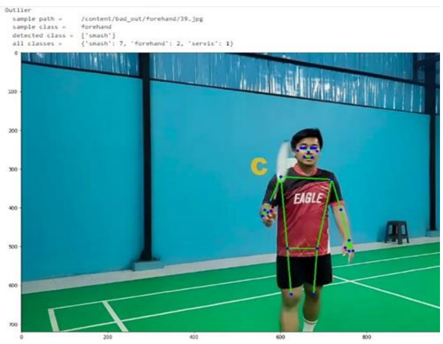

Figure 9. Example of data outliers

A second round of bootstrapping is performed to aggregate all the data from each distinct class into a single set once the outliers have been removed from the data. This information was produced in order to train the categorization mode.

4) *Training Data*

We categorize landmarks using a number of supervised learning classification models, such as logistic regression, random forest, and k-nearest neighbor. The results of the data training process are measured using a variety of common matrices, such as matrix accuracy, recall, and precession, in order to compare model performance.

a) *Confusion Matrix*

For 10% of the entire data, or 88 data, Figure 10 shows the positive results of the data test using the four logistic regression, k-nearest neighbor, and random forest classification models. The logistic regression model exhibits the most significant inaccuracy when compared to other classification models. In the Forehand Engineering class, 26 out of the 34 data were correctly predicted, with only 8 being incorrect. In the Service Engineering class, 18 out of the 21 data were correctly predicted, and just 2 were incorrect. Only three of the 33 data in the Smash Engineering class were incorrectly predicted; thirty of the 33 were correctly predicted. Based on classification from a random forest, the random forest model has the lowest error and maximum degree of relevance.

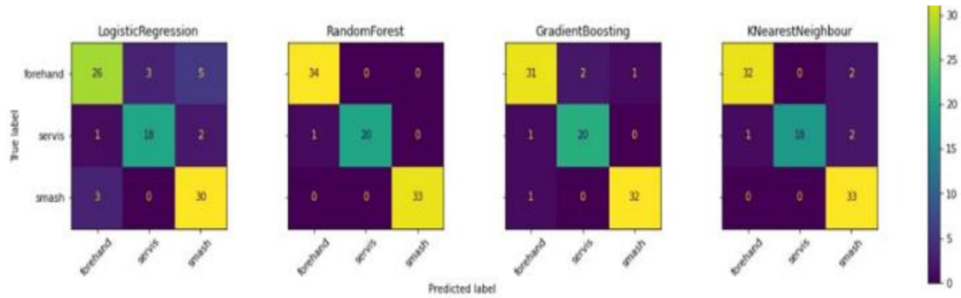

Figure 10. Confusion Matrix

b) Accuracy, Recall, and Precision

The model's accuracy reveals how successfully it can categorize data. In **Table 5**, the accuracy value from this study is displayed.

Table 5. Accuracy Model values

No	Algorithm Name	Accuracy
1	Logistics Regression	0.840
2	Random Forest	0.988
3	K-Nearest Neighbor	0.943

5) *Testing the Classification Model with New Data*

The data utilized in this testing approach is different from the data used in testing the previous model since it contains data in the form of videos or photographs. The random forest model was employed as a classifier in the testing that led to the model's selection. This goal was achieved in order to assess the model's real performance in categorizing badminton approaches as well as crucial spots and landmarks.

a) *Model Testing with Image Input*

Ten images from each category of badminton technique make up the 30 photos that make up the input data for this testing procedure. The confusion matrix table was made using the new data and the testing results, and it appears as follows

Table 6. Confusion Matrix

Confusion Matrix		**Predicted Class**		
		Forehand	***Service***	***Smash***
Actual Class	***Forehand***	7	3	0
	Service	2	8	0
	Smash	5	2	3

According to the aforementioned confusion matrix, the prediction findings using the new test data show relatively good results in a number of classes, including the Forehand Engineering class and the Smash Technique class. While each of these courses properly predicted seven out of ten pieces of data for Forehand Engineering, Smash Engineering only managed to predict two out of ten pieces of data. Service Engineering, with a forecasted rating of 8 out of 10, has the best accuracy outcomes. The following accuracy estimates apply to all anticipated courses

$$\text{accuracy} = \frac{\text{Total True Positive}}{\text{Total sampel}} \quad (1)$$

$$\text{Accuracy} = \frac{18}{30} \times 100 = 60\%$$

b) Model Testing with Video Input

This test design aimed to evaluate models that used video as an input for posture estimation, model inference, and classification. The exam's flow is divided into the following phases.

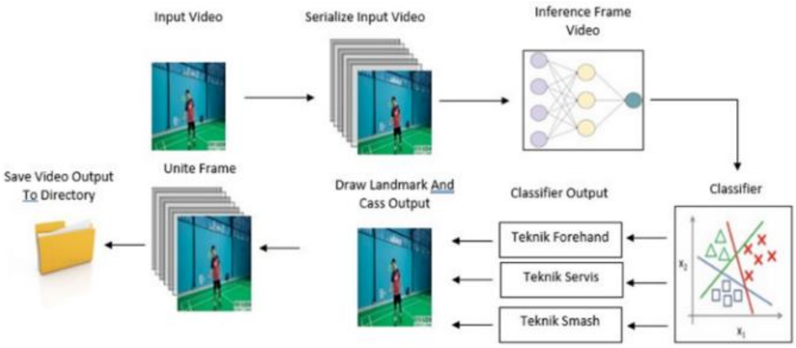

Figure 11. Video Inference Flowchart

The process begins with defining the input video, then serializing the video into frames, frame inference with Mediapipe, classifying the output landmarks using a classification model, drawing the landmark pose, prediction class, and prediction probability of video frames, and finally finishing with drawing the landmark pose. The frames are reassembled and stored in MOV format before being placed in the correct directory once all of the frames have been categorised and inferred. Here is an example of a picture produced via video inference:

Figure 12. Service Engineering video inference results

Figure 13. Forehand technique video inference results

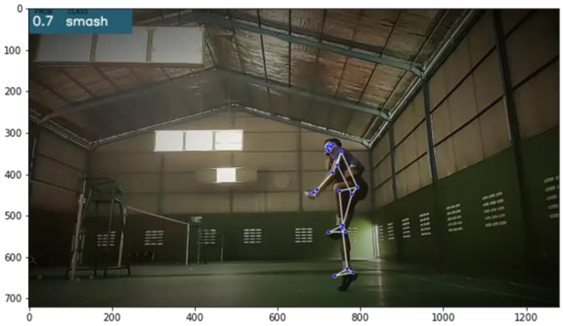

Figure 14. The Smash Technique video inference results

B. Discussion

The implications of these findings for badminton training and technique improvement are significant. By utilizing a CNN-based image recognition approach and pose recognition technologies such as BlazePose and Mediapipe Pose Solution, badminton coaches and players can gain deeper insights into the strategies used in the game. Accurate identification of poses in training videos can assist in evaluating player performance, identifying their strengths and weaknesses, and designing more effective and focused training programs.

Comparisons with existing methods or studies can highlight the unique contributions of the proposed approach in this research. While many previous studies have employed image recognition techniques and CNNs to analyze athlete movements in various sports, the use of BlazePose and Mediapipe Pose Solution in the context of badminton has been relatively unexplored. By leveraging these cutting-edge technologies, this research provides a novel contribution to the field of sports analytics, particularly in understanding and enhancing badminton gameplay strategies.

During the research, we encountered some unexpected challenges. One of them was the complexity of processing and managing a large dataset, especially in ensuring sufficient representation of various badminton techniques. Additionally, we faced difficulties in handling outliers or unusual poses in the dataset, requiring additional efforts in data preprocessing and model tuning.

Nevertheless, our findings indicate that the proposed approach can effectively address these challenges, yielding significant classification accuracy. This highlights the considerable potential of utilizing image recognition technology and CNNs to enhance understanding and performance in badminton.

Overall, this research demonstrates that the use of CNN-based image recognition approaches and pose recognition technologies can make valuable contributions to badminton training and analysis. The practical implications of these findings can assist coaches and players in improving their on-court performance, while their theoretical contributions pave the way for further research in the field of sports analytics.

Conclusion

This study has successfully demonstrated the effectiveness of using Convolutional Neural Network (CNN) in conjunction with BlazePose architecture and Mediapipe Pose Solution tools for classifying badminton strategies based on poses in training videos.

Key findings from this research include achieving significant classification accuracy ranging from 80% to 90% with the implemented CNN model. This indicates the robustness and reliability of the proposed approach in accurately categorizing various badminton techniques.

The implications of this research are significant, particularly in the realm of sports training and analysis. By accurately classifying badminton strategies, coaches and players can gain valuable insights into their performance, identify areas for improvement, and tailor training programs more effectively.

However, it is important to acknowledge some limitations of this study. One limitation is the challenge of collecting a large and diverse dataset to ensure comprehensive coverage of various badminton techniques and scenarios. Additionally, the model's performance may be impacted by outliers or unusual poses in the dataset, highlighting the need for further data preprocessing and refinement.

For future research, it is recommended to explore additional techniques to enhance the accuracy and efficiency of the classification model. This may include investigating advanced CNN architectures, incorporating temporal information from video sequences, and addressing the challenges associated with outlier poses. Furthermore, future studies could extend the application of this method to other sports disciplines, thereby contributing to the broader field of sports analytics and enhancing training methodologies.

In conclusion, this research provides a valuable contribution to the field of sports technology by presenting an effective approach for classifying badminton strategies based on pose recognition in training videos. The findings and methodologies presented here have the potential to benefit coaches, players, and sports enthusiasts alike, ultimately enhancing performance and promoting a deeper understanding of the game.

References

[1] Hendriyana and Y. H. Maulana, "Identifikasi Jenis Kayu menggunakan Convolutional Neural Network dengan Arsitektur Mobilenet," J. RESTI (Rekayasa Sist. dan Teknol. Informasi), vol. 4, no. 1, pp. 70–76, 2020, doi: 10.29207/resti.v4i1.1445.

[2] F. Endrianti, W. Setiawan, and Y. Wihardi, "Sistem Pencatatan Kehadiran Otomatis di Ruang Kelas Berbasis Pengenalan Wajah Menggunakan Metode Convolutional Neural Network (CNN)," JATIKOM - J. Apl. dan Teori. Ilmu Komput., vol. 1, no. 1, pp. 40–44, 2018, doi: 10.17509/jatikom.v1i1.25146.

[3] H. S. Ibrahim, M. Si, U. N. Wisesty, F. Informatika, and U. Telkom, "Analisis Deep Learning Untuk Mengenali Qrs Kompleks Pada Sinyal Ecg Dengan Metode CNN," e-Proceeding Eng., vol. 5, no. 2, pp. 3718–3725, 2018.

[4] M. Zufar and B. Setiyono, "Convolutional Neural Networks Untuk Pengenalan Wajah Secara Real-Time," J. Sains dan Seni ITS, vol. 5, no. 2, p. 128862, 2021.

[5] F. Han, J. Yao, H. Zhu, and C. Wang, "Underwater Image Processing and Object Detection Based on Deep CNN Method," J. Sensors, vol. 2020, p. 20, 2020, doi: 10.1155/2020/6707328.

[6] M. Elleuch, R. Maalej, and M. Kherallah, "A New design based-SVM of the CNN classifier architecture with dropout for offline Arabic handwritten recognition," Procedia Comput. Sci., vol. 80, pp. 1712–1723, 2016, doi: 10.1016/j.procs.2016.05.512.

[7] B. Sekeroglu and I. Ozsahin, "Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks," SLAS Technol., vol. 25, no. 6, pp. 553–565, 2020, doi: 10.1177/2472630320958376.

[8] P. Ma, "Recognition of handwritten digit using convolutional neural network," Proc. - 2020 Int. Conf. Comput. Data Sci. CDS 2020, vol. 19, no. 2, pp. 183–190, 2020.

[9] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," Procedia Comput. Sci., vol. 132, pp. 679–688, 2018, doi: 10.1016/j.procs.2018.05.069.

[10] P. Wang, X. Zhang, and Y. Hao, "A Method Combining CNN and ELM for Feature Extraction and Classification of SAR Image," J. Sensors, vol. 2019, doi: 10.1155/2019/6134610.

[11] A. Patil and M. Rane, "Convolutional Neural Networks: An Overview and Its Applications in Pattern Recognition," Smart Innov. Syst. Technol., vol. 195, pp. 21–30, 2021, doi: 10.1007/978-981-15-7078-0_3.

[12] W. Lee, K. Ko, Z. Geem, K. Sim, and E. Engineering, "Method that Determining the Hyperparameter of CNN using HS algorithm," J. Korean Inst. Intell. Syst., vol. 27, no. 1, pp. 22–28, 2017.

[13] M. Sadeghi et al., "PERSIANN-CNN: Precipitation estimation from remotely sensed information using artificial neural networks–convolutional neural networks," J. Hydrometeorol., vol. 20, no. 12, pp. 2273–2289, 2019, doi: 10.1175/JHM-D-19-0110.1.

[14] Y. Supardi and D. Dede, Semua Bisa Menjadi Programmer Python Case Study. Jakarta: Elex Media Komputindo, 2020.

[15] A. R. Syulistyo, D. M. J. Purnomo, M. F. Rachmadi, and A. Wibowo, "Convolutions Subsampling Convolutions Gaussian connection Full connection Full connection Subsampling," JIKI (Jurnal Ilmu Komput. dan Informasi) UI, vol. 9, no. 1, pp. 52–58, 2020.

[16] P. A. Nugroho, I. Fenriana, and R. Arijanto, "Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) Pada Ekspresi Manusia," Algor, vol. 2, pp. 12–21, 2020.

[17] P. Lakhani, B. Sundaram, “Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks”. Radiology 284:574–582, 2021, doi: 10.1148/radiol.2017162326.

[18] F. Lu, F. Wu, P. Hu, Z. Peng, D. Kong D, "Automatic 3D liver location and segmentation via convolutional neural network and graph cut". Int J Comput Assist Radiol Surg 12:171–182, 2022, doi: 10.1007/s11548-016-1467-3.

[19] F. Milletari, N. Navab, S-A. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image

segmentation". In: Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV) , 2021, doi: 10.1109/3DV. 2016.79.

[20] PF. Christ, MEA. Elshaer, F. Ettlinger et al, "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields". In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W (eds) Proceedings of Medical image computing and computer-assisted intervention –MICCAI 2021, doi: 10.1007/978-3-319-46723-8_48.

[21] Shafira, Tiara. (2018). Implementasi Convolutional Neural Networks Untuk Klasifikasi Citra Tomat Menggunakan Keras. Skripsi. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Islam Indonesia : Yogyakarta.

[22] Dewa C. K., Fadhilah A. L. and Afiahayati. (2018). Convolutional Neural Networks for Handwritten Javanese Character Recognition IJCCS (Indonesian Journal of Computing and Cybernetics Systems) 93-94, doi: 10.22146/ijccs.31144

[23] F. Han, J. Yao, H. Zhu, and C. Wang, “Underwater Image Processing and Object Detection Based on Deep CNN Method,” J. Sensors, vol. 2020, p. 20, 2020, doi: 10.1155/2020/6707328.