Detection of similarities and differences within the same shot movement using artificial intelligence-based performance analysis: An example of a tennis service

Detección de similitudes y diferencias dentro de un mismo movimiento de golpeo mediante un análisis del rendimiento basado en inteligencia artificial: ejemplo del servicio en tenis

Takashi Jindo[1] *, Yusuke Satonaka[2], Ryosuke Wakamoto[2], Michitaka Iida[2], Hikari Suzuki[3], Hirotaka Shiraishi[3] and Daisuke Mitsuhashi[4]

1 Division of Art, Music, and Physical Education, Osaka Kyoiku University, 4-698-1 Asahigaoka, Kashiwara, Osaka 582-8582, Japan.
2 Information Services International-Dentsu, LTD, 2-17-1 Konan, Minato-ku, Tokyo 108-0075, Japan
3 Master's Program in Physical Education, Health and Sport Sciences University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8574, Japan.
4 Faculty of Health and Sport Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8574, Japan.

Received: 20-06-2023

Accepted: 07-10-2023

Abstract

Artificial intelligence (AI) -based performance analysis has the potential to support feedback in coaching; however, a useful method has not yet been proposed. This study aims to develop an AI-based performance analysis to support tennis coaching. Specifically, we investigate the accuracy of detecting similarities and differences within the same shot movement. The participants were two tennis players with more than ten years of tennis experience at the regional level. This study targeted service in tennis and videos of the 1st and 2nd service from both sides (number of services: 40 attempts) were recorded using a smartphone located on the fence behind the participant. The analysis code was executed in Python, and the main part involved the use of BlazePose, which estimates the X-, Y-, and Z-coordinates of a human pose. Video clips of 2 s were cut, with a 1 s overlap between each clip, and one of the clips was manually chosen as the standard clip. The clips were compared with the comparison clips, and the difference scores for the total and each body part were automatically calculated. An AI-based analysis was conducted considering 12 conditions combining the 1st and 2nd services from both sides and different players. As a result, a certain accuracy (≥ 70%) was confirmed for detecting overlapping phases between clips. Moreover, manually evaluated body parts that showed different movements by a certified coach corresponded to the top three different parts in the AI-based analysis for 8 of the 12 conditions. The proposed AI-based performance analysis can effectively extract similar or overlapping phases and suggest body parts exhibiting different movements.

Keywords: *Performance analysis, motion analysis, artificial intelligence (AI), tennis, service.*

Corresponding author: Takashi Jindo, jindo-t93@cc.osaka-kyoiku.ac.jp

Cite this article as:
Jindo, T., Satonaka, Y., Wakamoto, R., Iida, M., Suzuki, H., Shiraishi, H., & Mitsuhashi, D. (2023). Detection of similarities and differences within the same shot movement using artificial intelligence-based performance analysis: An example of a tennis service. *International Journal of Racket Sports Science, 5*(1), xxxx.

Resumen

El análisis del rendimiento basado en inteligencia artificial (IA) tiene el potencial de apoyar la retroalimentación en el entrenamiento. Sin embargo, aún no se ha propuesto un método útil. El objetivo de este estudio es desarrollar un análisis del rendimiento basado en IA para apoyar el entrenamiento de tenis. En concreto, se investiga la precisión en la detección de similitudes y diferencias dentro de un mismo movimiento de golpeo. Los participantes fueron dos tenistas con más de diez años de experiencia en tenis a nivel regional. Este estudio se centró en el servicio en tenis y se grabaron videos de los dos primeros servicios desde ambos lados de la cancha (número de servicios: 40 intentos) con un teléfono inteligente situado en la valla detrás del participante. El código de análisis se ejecutó en Python, y la parte principal involucró el uso de BlazePose, que estima las coordenadas X, Y y Z de una posición humana. Se cortaron videos de 2 s, con un solapamiento de 1 s entre cada video, y se eligió manualmente uno de ellos como el video estándar. Los videos se compararon con los de comparación y se calcularon automáticamente las puntuaciones de diferencia para el total y para cada parte del cuerpo. Se realizó un análisis basado en IA que consideraba 12 condiciones y combinaba los dos primeros servicios desde ambos lados y de los diferentes jugadores. Como resultado, se confirmó cierta precisión (≥ 70%) en la detección de fases solapadas entre videos. Además, las partes del cuerpo evaluadas manualmente que mostraban movimientos diferentes por un entrenador certificado correspondían con las tres primeras partes diferentes del análisis basado en IA para 8 de las 12 condiciones. El análisis de rendimiento basado en IA propuesto puede extraer eficazmente fases similares o solapadas y sugerir partes del cuerpo que muestran movimientos diferentes.

Palabras clave: *Análisis del rendimiento, análisis del movimiento, inteligencia artificial (IA), tenis, servicio.*

INTRODUCTION

Sports performance analyses using objective information are increasingly being conducted. In tennis, the target of this study, numerical data such as the success rates of various shots and rallies are often used as objective information (O'Donoghue, 2005). However, a performance analysis should adopt qualitative data, such as videos, to obtain detailed information. Qualitative data analysis takes the form of inputting specific events and labels using specialized software. Although this method can be useful for understanding the characteristics of each scene of play, the interpretation of the information obtained can be influenced by the experience and subjectivity of the players and coaches.

A method for automatically analyzing video data was developed based on the object-detection technology of artificial intelligence (AI), such as machine learning and its division deep learning (Brady et al., 2021; Cust et al., 2019). A systematic review stated that machine learning has been increasingly adopted in tennis tracking or analyzing both player and ball movements (Takahashi et al., 2022). More models are expected to feature deep learning owing to the development of better hardware and advantages of achieving more efficient model learning on large data inputs (Cust et al., 2019), which are suitable for racket sports because of the large number of attempted shots and movements. A systematic review of this research field (Cust et al., 2019) summarized inertial motion unit (IMU)-based and vision-based AI-based performance analyses for various sports. Different methods may be appropriate for different types of sports or situations (practice session or match). For example, in tennis, the vision-based method would be more useful because there is no need to attach devices to the player and there is potential for application in real matches. The systematic review (Cust et al., 2019) has summarized three studies that used vision-based analysis for tennis (Ó Conaire et al., 2010; Shah et al., 2007; Zhu et al., 2006). However, these studies only classified shot types such as service, forehand, or backhand strokes. Following a systematic review (Cust et al., 2019), Cai et al. (2020) conducted a more detailed AI-based analysis on 12 shots in tennis, and confirmed that there was significant confusion within the same shot, such as between the kick service and slice service.

Human pose estimation has significantly advanced. BlazePose, a lightweight convolutional neural network architecture for human pose estimation developed by Google Research (Bazarevsky et al., 2020) might be useful for tennis performance analysis. BlazePose has the ability to estimate the X-, Y-, and Z-coordinates. Therefore, this method is suitable for tennis analysis, which requires depth estimation for body direction or stance during shot movements. To date, shot classification for six-shot movements in cricket (Devanandan et al., 2021) and estimation of multiple joint angles during tennis service (Yoshida et al., 2021) using BlazePose have been reported. Although these studies suggest the application possibilities of BlazePose in racquet sports, an accumulation of study findings is required to support sports coaching. Information on similarities or differences within the same shot is useful in sports coaching; therefore, AI-based analysis is expected to play a supportive role.

This study aimed to develop an AI-based performance analysis for tennis coaching applications. Specifically, we investigate the accuracy of detecting similarities and differences within the same shot movement.

MATERIAL AND METHODS

Study design and participant

An experimental study was conducted in a tennis court. The tennis court was blue and light blue in color, and the surface was hard. The study participants were two tennis players, who were also the authors of this study. The participants are experienced tennis players with more than ten years of experience at a regional level. The self-reported height and weight of the two players were as follows: 173 cm, 62 kg for Player 1, and 170 cm, 68 kg for Player 2. This study was approved by the Ethics Committee of University of Tsukuba (approval number: Tai 022-80 and Tai 022-80-1).

Data preparation

This study focuses on tennis service as a target for performance analysis. The reason for focusing on service is that the first service points have a large impact on the win or loss of a match for both professional and junior players (Kovalchik & Reid, 2017). Thus, service is considered to have a high priority in improving tennis shot skills. In addition, the service would be a suitable target for the first step of the developed analysis because it can be performed by one player and does not need to consider the inclusion of opponents in the video.

The proposed method required two target videos. The first was a standard video that included one or more attempts at the target shot. A standard video is assumed to include an ideal movement or movement prior to specific training implementation. The second video was used for comparison. The comparison video included multiple attempts of the shot for comparison. In this study, some service videos of the two study participants were recorded, and certain conditions were set to evaluate the developed method (details are described below).

The videos were recorded at 1920 × 1080 pixels and 60 frame per second (fps) using an iPhone (Apple Inc.). The iPhone was located approximately 1.7 m high on the fence 8.5 m behind the participant and was the same for all conditions. The fence was stable, and leaning over did not affect the camera angle.

Development of AI-based performance analysis

This study developed an AI-based performance analysis method that uses tennis videos to automatically detect the differences and similarities within the same shot movement. The analysis was conducted using Google Colaboratory, an online execution environment for Python (https://www.python.org/). Target videos were uploaded to Google Drive's cloud storage system and imported into the environment. The programming code was prototyped by researchers from Information Services International-Dentsu, LTD. The code was modified based on a preliminary study and discussions between the company and university researchers. Subsequently, the video data were gathered and analyzed by university researchers.

An overview of the AI-based performance analysis is shown in Figure 1. We adopted BlazePose (Bazarevsky et al., 2020), which is a lightweight convolutional neural network architecture for human pose estimation, for the main part of the analysis. BlazePose is one of the models in the Mediapipe framework developed by Google that offers customizable machine learning solutions for processing multimodal data. The technology is open source and available to the public. Although BlazePoze can estimate the coordinates of 33 body parts, the developed method targeted 13 body parts (nose, shoulders, elbows, wrists, left and right hips, knees, and ankles) that are important for tennis shot movements. The nose was included because its position would be useful for ascertaining the status of neck rotation, extension, and flexion.

The recorded video was converted from the MOV file to a GIF file and analyzed using developed functions, including BlazePose. In addition to the MOV file, the MP4 video format is also applicable for the analysis. In addition to converting to a GIF file, the GIF file was clipped to short-duration clips with an overlapping duration to conduct an analysis targeting the appropriate phase, such as before and after the impact. A phase indicates the entire or partial movement of a shot, whereas the entire phase indicates a shot. In this study, the clipped duration was set to 2 s, with a 1 s overlap between each clip. One of the clips that included the impact of the ball and racket was manually selected as the standard clip. A clip duration of 2 s mostly covered the entire service phase from toss-up to after impact.

These functions transform the skeletal coordinate information while excluding differences in the recording angle (i.e., between deuce and advantage sides) or position of the camera (i.e., this was not applicable to this study because of the same camera position). Specifically, a skeletal part (the left shoulder in this study) was set as the origin of the coordinates, and parallel shifts and rotations of all spatial coordinates were performed with respect to the origin. The skeletal coordinate information extracted by the functions was used to calculate the difference score between the videos.

To calculate the difference scores, the Dynamic Time Warping algorithm (Sakoe & Chiba, 1978; https://tslearn.readthedocs.io/en/stable/user_guide/dtw.html) was used to consider the movement speed. Scores were calculated for each body part, with lower scores indicating smaller differences between the standard and comparison clips. The average of all body part scores was considered the total difference score for movement in this study. The scores ranged from zero to no upper limit.

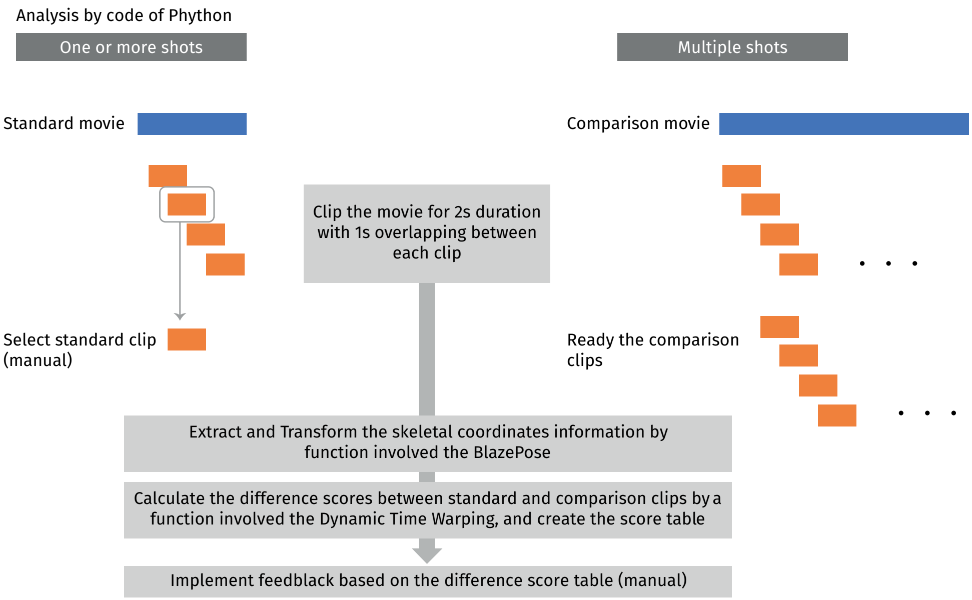

Figure 1. Overview of AI-based analysis.

These results were summarized in a score table that was sorted by the lowest total score and contained the scores of each body part. The score table is automatically output, and the results are interpreted by users.

The developed code can be demonstrated here (https://drive.google.com/drive/folders/11UG48uyxXfjIUTE6v5YGk9eZ109WRPdU?usp=sharing).

Evaluation

To achieve this goal, we conducted two evaluations. The first was to evaluate whether similar phase clips (i.e., from toss-up to immediately after the impact) were in the upper part of the score table to clarify whether the AI-based analysis extracted the overlapping phase of the standard clip from the comparison clips. To conduct the evaluation, the following conditions for service videos from the two study participants were gathered:1st and 2nd services for five shots on the deuce and advantage sides, respectively. The collected videos were used to conduct an AI-based analysis of comparisons within the same condition, between different conditions, and between different players; 12 conditions were conducted. The accuracy (%) was calculated by referring to previous studies (Cai et al., 2020; O Conaire et al., 2010). Notably, the top 12 clips within the same condition and top 15 clips between different conditions and players from the score table were considered for the calculation. Specifically, the accuracy was calculated by dividing the number of correct detections (i.e., overlapping the phase) by the total number of compared files. The top clips were considered because clipped comparison movies were assumed to contain at least 12 or 15 overlapping phase clips with the standard clip. The 2 s standard clip mostly contained the entire phase of service whilst the comparison movies contained four shots in the case of the same condition, and five shots in the case of different conditions and players. Therefore, although the number of clips for a single shot is dependent on the timing of clipping, a single shot can be clipped for three clips with 1 s overlap. Accordingly, the number of targeting clips was calculated by multiplying three movies by four or five shots. In addition, Spearman correlation coefficients (ρ) between correct detection and difference scores were calculated to understand these relationships.

The second objective was to evaluate the correspondence between manual and AI-based analyses for the detection of different movements in each body part. The target clips for the evaluation were set as standard clips, and one comparison clip showed the lowest total score in the first evaluation under each condition. When the clip did not have an overlapping phase between clips, the clip with the second-lowest total score was used. Manual evaluation was conducted by one of the authors who is a qualified Japan Sports Association Instructor of Tennis. The evaluator observed and compared the clips from the perspective of different movements and described applicable body parts that had different movements from the standard clip. As the left shoulder was set as the origin, this part was excluded from the evaluation. The evaluation was conducted under concealed conditions based on the results of the AI-based analysis. After completion of the manual evaluation, body parts with any description were compared with the scores and ranks of the AI-based analysis.

The evaluation of shot movement was conducted using the open-source movement analysis software Kinovia (https://www.kinovea.org/) and its functions, such as two playback screens, slow motion, and rewinding. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp. was used for data analysis. The level of statistical significance was set at $P < 0.05$.

RESULTS

Table 1 shows the accuracy of extracting the overlapping phase of the standard clip from the comparison clips. The duration of the target videos ranged from 34 to 45 s; thus, 34 to 45 clips were created. The average accuracy when comparing within the same condition was 79.2 % for both deuce and advantage sides. The average accuracy when comparing between different conditions were as follow: 73.3 % in case of side comparison, 70.0 % in case of the first and second service comparisons in both deuce and advantage sides. Among these, the accuracy in case of the side comparison for Player 2 was the lowest (46.7 %.), whilst that between different players was 80.0 and 40.0 % for the deuce and advantage sides, respectively.

Table 2 shows the correlation between the correct detection and difference scores. The total and difference scores for each body part mostly showed a negative correlation with correct detection, which indicates that clips with lower difference scores are more likely to overlap. For the overall correlations, all difference scores showed a statistically significant correlation with correct detections. Moreover, side comparison for Player 2, which showed low accuracy of 46.7 %, had statistically positive correlation in the nose (ρ = 0.35), left elbow (ρ = 0.48), and left wrist (ρ = 0.40). For the different players' comparison in the advantage side (accuracy: 40.0 %), some body parts showed statistically significant negative correlations (left elbow: ρ = -0.36 and left wrist: ρ = -0.35) while five body parts showed positive correlations.

Table 3 shows a comparison of the results between the manual and AI-based analyses for comparison within the same conditions. For the deuce side of Player 2, the manually evaluated body parts that showed different movements corresponded to the top three different parts in the AI-based analysis. However, for the deuce side for Player 1 and the advantage side for Player 2, manually evaluating different movements of the body parts had the lowest difference scores in the AI-based analysis (considered to be similar movements). For the advantage side of Player 1, there were no confirmed body parts that had different movements in the manual evaluation, whereas the left elbow showed the highest difference score.

Table 4 shows a comparison of the results of the manual and AI-based analyses for different conditions. One of the manually evaluated body parts that exhibited different movements corresponded to the top three different parts in the AI-based analysis under all conditions.

Table 5 shows a comparison of the results of the manual and AI-based analyses for different players. On the source side, two of the manually evaluated body parts that showed different movements corresponded to the top two or three parts in the AI-based analysis. For the advantage side, manually evaluated different movements of body parts were the top four difference scores in the AI-based analysis.

Table 1.
Accuracy for extracting overlapped phase of the standard clip from the comparison clips.

Condition		Player 1		Player 2		Total	
Standard	Comparison	Number of correct detections	Accuracy (%)	Number of correct detections	Accuracy (%)	Number of correct detections	Accuracy (%)
Comparison within same condition							
1st service, deuce side		11	91.7	8	66.7	19	79.2
1st service, advantage side		11	91.7	8	66.7	19	79.2
Comparison within different conditions							
1st service, deuce side	1st service, advantage side	15	100.0	7	46.7	22	73.3
1st service, deuce side	2nd service, deuce side	12	80.0	9	60.0	21	70.0
1st service, advantage side	2nd service, advantage side	10	66.7	11	73.3	21	70.0
Comparison within different players							
1st service, deuce side	-	-	-	-	-	12	80.0
1st service, advantage side	-	-	-	-	-	6	40.0

The top 12 or 15 similar comparison clips were targeted for each condition.
Accuracy (%) = number of correct detections / total number of compared files.

Table 2.
Correlation between correct detection and difference scores.

Condition		Player	Difference score																										
			Total		Nose		Right shoulder		Left shoulder	Right elbow		Left elbow		Right wrist		Left wrist		Right hip		Left hip		Right knee		Left knee		Right ankle		Left ankle	
Standard	Comparison		ρ	P-value	ρ	P-value	ρ	P-value	(reference)	ρ	P-value	ρ	P-value	ρ	P-value	ρ	P-value	ρ	P-value	ρ	P-value	ρ	P-value	ρ	P-value	ρ	P-value	ρ	P-value
Comparison within same condition																													
1st service, deuce side	-	1	**-0.42**	**0.01**	**-0.47**	**0.00**	-0.15	0.37		-0.19	0.26	**-0.59**	**0.00**	-0.01	0.97	**-0.56**	**0.00**	**-0.43**	**0.01**	-0.28	0.09	**-0.49**	**0.00**	-0.18	0.28	**-0.45**	**0.01**	-0.19	0.26
		2	-0.21	0.17	-0.01	0.95	0.01	0.93		-0.05	0.72	0.12	0.43	0.00	0.98	0.06	0.70	-0.07	0.65	0.00	1.00	**-0.49**	**0.00**	**-0.44**	**0.00**	**-0.39**	**0.01**	**-0.40**	**0.01**
1st service, advantage side	-	1	**-0.54**	**0.00**	-0.30	0.08	-0.28	0.11		**-0.57**	**0.00**	**-0.40**	**0.02**	**-0.44**	**0.01**	**-0.39**	**0.02**	-0.16	0.37	-0.25	0.15	-0.34	0.05	-0.29	0.10	-0.34	0.05	**-0.37**	**0.03**
		2	-0.01	0.94	0.01	0.95	-0.11	0.47		-0.17	0.26	-0.28	0.06	-0.24	0.12	-0.24	0.11	-0.06	0.72	0.04	0.79	0.06	0.60	0.12	0.45	0.14	0.37	0.25	0.11
Comparison between different conditions																													
1st service, deuce side	1st service, advantage side	1	**-0.59**	**0.00**	-0.32	0.06	-0.29	0.09		-0.26	0.14	-0.31	0.07	**-0.35**	**0.04**	-0.23	0.18	**-0.52**	**0.00**	**-0.43**	**0.01**	**-0.46**	**0.01**	**-0.46**	**0.01**	**-0.45**	**0.01**	-0.24	0.17
		2	0.10	0.53	**0.35**	**0.02**	0.14	0.37		0.01	0.95	**0.43**	**0.00**	0.19	0.22	**0.40**	**0.01**	-0.10	0.52	-0.05	0.72	-0.20	0.20	-0.22	0.14	-0.17	0.26	-0.28	0.06
1st service, deuce side	2nd service, deuce side	1	**-0.40**	**0.02**	**-0.44**	**0.01**	-0.14	0.44		-0.27	0.12	-0.09	0.63	**-0.40**	**0.02**	-0.19	0.28	**-0.45**	**0.01**	**-0.34**	**0.05**	**-0.45**	**0.01**	-0.14	0.44	-0.32	0.06	-0.07	0.68
		2	-0.10	0.51	-0.03	0.87	-0.20	0.20		-0.08	0.59	-0.27	0.09	-0.18	0.25	-0.15	0.35	0.15	0.34	0.24	0.13	-0.19	0.22	-0.05	0.78	-0.20	0.20	-0.06	0.70
1st service, advantage side	2nd service, advantage side	1	-0.23	0.17	-0.29	0.09	-0.16	0.36		**-0.37**	**0.03**	-0.09	0.60	-0.24	0.15	-0.16	0.36	-0.02	0.93	-0.08	0.64	-0.10	0.57	-0.26	0.13	-0.01	0.98	-0.08	0.64
		2	-0.08	0.59	-0.14	0.35	**-0.33**	**0.03**		-0.09	0.56	-0.01	0.96	-0.01	0.93	-0.06	0.60	-0.09	0.56	-0.05	0.76	-0.09	0.55	-0.17	0.27	-0.10	0.53	-0.24	0.12
Comparison between different players																													
1st service, deuce side	-	1	-0.27	0.07	**-0.43**	**0.00**	-0.11	0.48		0.04	0.80	-0.15	0.31	0.02	0.91	-0.29	0.05	0.00	0.97	-0.24	0.10	-0.19	0.22	-0.28	0.06	-0.08	0.60	-0.15	0.31
1st service, advantage side	-	2	-0.12	0.45	0.14	0.35	-0.07	0.65		0.07	0.63	**-0.36**	0.02	-0.19	0.21	**-0.35**	0.02	0.08	0.58	0.15	0.32	0.03	0.86	-0.03	0.86	-0.16	0.30	-0.21	0.17
Overall			**-0.20**	**0.00**	**-0.12**	**0.01**	**-0.12**	**0.01**		**-0.16**	**0.00**	**-0.16**	**0.00**	**-0.13**	**0.00**	**-0.18**	**0.00**	**-0.12**	**0.01**	**-0.09**	**0.05**	**-0.20**	**0.00**	**-0.18**	**0.00**	**-0.17**	**0.00**	**-0.15**	**0.00**

Table 3.
Comparison of the results between manual and AI-based analysis for the comparison within the same condition.

Items		Body parts												Note
		Nose	Right shoulder	Left shoulder	Right elbow	Left elbow	Right wrist	Left wrist	Right hip	Left hip	Right knee	Left knee	Right ankle	Left ankle	
1st service, deuce side, Player 1															
Manual analysis				Reference								Deeply bent when toss-up			
AI-based analysis	Difference score	13.17	14.08		13.24	12.79	13.09	12.4	13.21	12.16	12.28	11.1	12.07	11.41	
	Rank of difference	4	1		2	6	5	7	3	9	8	12	10	11	
1st service, deuce side, Player 2															
Manual analysis				Reference						Down at the impact (whole body tilted to the left)					
AI-based analysis	Difference score	14.26	13.14		13.89	14.22	13.96	14.68	14.8	14.54	14.06	13.71	13.76	13.58	
	Rank of difference	4	12		8	5	7	2	1	3	6	10	9	11	
1st service, advantage side, Player 1															
Manual analysis				Reference											Limited overlapped phase (only for toss-up), no different parts observed manually
AI-based analysis	Difference score	14.89	14.04		13.74	15.16	13.67	14.88	13.87	13.61	14.14	14.13	14.91	15.05	
	Rank of difference	4	8		10	1	11	5	9	12	6	7	3	2	
1st service, advantage side, Player 2															
Manual analysis				Reference	Located slightly on the outside after the impact (the side is open)										
AI-based analysis	Difference score	12.75	14.07		12.72	12.82	12.85	12.91	13.05	13.51	12.87	13.02	13.04	13.41	
	Rank of difference	11	1		12	10	9	7	4	2	8	6	5	3	

Table 4.
Comparison of the results between manual and AI-based analysis for the comparison between different conditions.

Items		Body parts													Note
		Nose	Right shoulder	Left shoulder	Right elbow	Left elbow	Right wrist	Left wrist	Right hip	Left hip	Right knee	Left knee	Right ankle	Left ankle	
1st service, deuce side - advantage side, Player 1															
Manual analysis			Closer to parallel to the net when the toss-up (also the direction of whole body)	Reference											
AI-based analysis	Difference score	14	14.49		13.32	14.12	13.39	14.04	13.41	13.44	13.63	13.09	13.97	13.64	
	Rank of difference	4	1		11	2	10	3	9	8	7	12	5	6	
1st service, deuce side - advantage side, Player 2															
Manual analysis				Reference			Large follow-through on landing after the impact		Located to the net side after the impact						Limited overlapped phase (only for follow-through)
AI-based analysis	Difference score	15.18	15.68		15.82	14.95	15.27	14.96	14.11	13.55	13.48	13.39	13.58	13.84	
	Rank of difference	4	2		1	6	3	5	7	10	11	12	9	8	
1st - 2nd service, deuce side, Player 1															
Manual analysis			More perpendicular to the ground just before the impact (more shoulder abduction)	Reference	Swing more to the upper right immediately after the impact								Elevated higher position on landing after the impact		
AI-based analysis	Difference score	13.72	12.82		13.69	13.52	13.59	12.88	13.94	13.58	12.87	12.62	12.98	12.29	
	Rank of difference	2	10		3	6	4	8	1	5	9	11	7	12	
1st - 2nd service, deuce side, Player 2															
Manual analysis				Reference				Large swing-up on landing after the impact	The pelvis rotates to the net side at an earlier time just before the impact				The pelvis is abducted and positioned more outward at the impact		
AI-based analysis	Difference score	13.33	13.48		15.11	14.09	15.21	14.29	14.04	13.76	14.15	13.71	14.05	13.97	
	Rank of difference	12	11		2	5	1	3	7	9	4	10	6	8	
1st - 2nd service, advantage side, Player 1															
Manual analysis			Almost parallel to the net when toss-up (and overall body orientation as well)	Reference		Bended on landing after the impact			Slow pelvic rotation to the net side from just before the impact to its end						
AI-based analysis	Difference score	13.31	12.76		13.48	13.79	13.42	14	13.38	13.25	13.22	13.4	13.3	13.56	
	Rank of difference	8	12		4	2	5	1	7	10	11	6	9	3	
1st - 2nd service, advantage side, Player 2															
Manual analysis				Reference					Slow pelvic rotation to the net side from just before the impact to its end				The pelvis is abducted and positioned more outward at the impact		
AI-based analysis	Difference score	12.48	12.5		12.21	12.55	12.74	12.72	12.77	12.58	13.43	13.87	13.88	14.07	
	Rank of difference	11	10		12	9	6	7	5	8	4	3	2	1	

Table 5.
Comparison of the results between manual and AI-based analysis for the comparison between different players

Items		**Body parts**													
		Nose	**Right shoulder**	**Left shoulder**	**Right elbow**	**Left elbow**	**Right wrist**	**Left wrist**	**Right hip**	**Left hip**	**Right knee**	**Left knee**	**Right ankle**	**Left ankle**	**Note**
1st service, deuce side, Player 1-2															
Manual analysis				Reference		Not swung up on landing after the impact (shoulder not abducted)		More perpendicular to the ground when loss-up (more shoulder abduction)	The pelvis rotates to the net side early just before the impact				The pelvis is abducted and positioned more outward at the impact		0
AI-based analysis	Difference score	13.87	14.78		14.39	14.3	14.48	14.51	14.72	13.99	13.18	12.84	13.41	12.99	
	Rank of difference	8	1		5	6	4	3	2	7	10	12	9	11	
1st service, advantage side, Player 1-2															
Manual analysis		Looking up when loss-up (neck is extended)		Reference											Limited overlapped phase (only for loss-up)
AI-based analysis	Difference score	15.1	14.13		14.61	13.26	14.03	13.4	15.2	15.48	15.45	14.68	13.8	13.75	
	Rank of difference	4	7		6	12	8	11	3	1	2	5	9	10	

DISCUSSION

This study developed an AI-based performance analysis for tennis coaching and investigated its accuracy in detecting similarities and differences within the same shot movement. As a result, a certain accuracy (≥70 %) was confirmed for detecting overlapping phases between clips. Moreover, the manually evaluated body parts that showed different movements corresponded favorably to the results of the AI-based analysis. Based on these results, the developed analysis can play a supportive role in finding observation points in tennis coaching.

Previous studies (Cai et al., 2020; Ó Conaire et al., 2010; Shah et al., 2007; Zhu et al., 2006) have mostly focused on shot classification (service, forehand, or backhand stroke). This study was conducted from an original perspective that attempted to extract similar phases. Results show that the accuracy was as high as 79.2 % when comparing within the same conditions, in the 70.0 - 73.3 % range when comparing between different conditions, and 80.0 % when comparing between different players in the deuce side. In a recent study (Cai et al., 2020), it was reported that a machine learning-based analysis was able to recognize 12 tennis actions with an accuracy of 62 %. Accordingly, most of the conditions in this study showed equivalent or higher accuracy than the previous study. The developed AI-based analysis is expected to provide effective feedback in coaching by automatically extracting overlapping phases with a certain accuracy.

However, low accuracy was confirmed under some conditions. Overall, accuracy was low for the conditions targeting Player 2. The lowest accuracy was observed in the side comparison for Player 2 (46.7 %). As the correlation with correct detection in the left elbow and wrist were mostly negative coefficients, except for the comparison within the same condition for Player 2, weighting these body parts for total difference scores would be effective in improving accuracy. Actually, when doubled the weight of difference scores in the left elbow and wrist for total score, the accuracy was 53.3%. The side comparison for Player 2 also showed low accuracy. It is possible that large differences in shot movements or positions between the sides affect the accuracy of extracting the overlapped phase. Adjusting or shortening the duration of clipping would be effective in improving accuracy. We attempted the solution for the side comparison for Player 2 by adjusting the clipping duration to 1 s with 0.5 s overlap. As the result, contrary to our expectations, the accuracy decreases to 6.6 %. Nevertheless, the adjustment of the clipping duration is also related to the phase length that the players and coaches want to focus on; thus, discretion of the user could be reflected. We assumed that the inclusion of extra movement after shots in the clips might have reduced accuracy, but this speculation was not applicable in this case. Another reason for the low accuracy could be that Player 2 wore a similarly colored shirt, and the estimator of the AI-based analysis was not able to estimate the human pose by changing the side and recording angle. Further investigation is required to determine the conditions that lead to low accuracy in the analysis.

This study attempts to extract the differences within the same shot movement using AI-based analysis. A previous study performed a comparison within the same shot in tennis (Cai et al., 2020), however, it reported that there was considerable confusion regarding the accuracy of the comparison. This study attempted to extract information on the differences in the movements of each body part that might be required for coaching. As a result, manually evaluated body parts that showed different movements by a certified coach corresponded to the top three different parts in the AI-based analysis for 8 of the 12 conditions. Although manual interpretation by users, such as players and coaches, automatically provides information on which body parts show different movements, it would help in effective and accurate motion analysis. However, some conditions showed contradictory results between manual and AI-based analyses. In this regard, the comparison within the same condition on the deuce side for Player 1 and the advantage side for Player 2 had the highest difference score for the right shoulder in the AI-based analysis. This may be because the internal and external rotation of the shoulder during service is fast, and the AI-based analysis would be judged as significantly different because of the video recording setting. In any case, no significant difference was observed between the clips because the services were performed consecutively in the comparison under the same conditions. It is expected that AI-based analysis would be useful for comparison in conditions that would have large differences, such as between early and late stages in a match or before and after a specific practice. Practical investigation of the application is needed. If rallies are targeted in the future, guards should be installed to protect cameras from ball hits. In addition, different scores in some clips with limited overlapping phases might have affected the movement included only in either clip. This may be an issue in practical applications. The proposed AI-based analysis should be treated as a supplemental tool to help coaching because selecting the reference video and interpreting the differences in movements are required to be conducted manually by the players or coaches. Even if this AI-based analysis underlines the differences in movement, the coach's assistance is still necessary if the players themselves have difficulty interpreting these differences.

This study had several limitations. First, this study conducted video recordings under identical conditions (participants' wear, camera location, type and setting, weather, court color, background, etc.). Therefore, further investigation is required to determine whether these conditions affect the analysis accuracy. Second, although variations in accuracy were observed in this study, this should be clarified by focusing on a variety of participants. Our aim was to compare services in various situations and investigate the perceptions of players and coaches to develop a useful AI-based analysis.

CONCLUSIONS

This study attempted to develop an AI-based analysis that plays a supportive role in tennis coaching, and investigated its accuracy-targeting services. Consequently, a certain accuracy in detecting similarities and differences between movements was confirmed. Although there were some issues that needed to be solved, this AI analysis could effectively extract similar or overlapping phases and suggest body parts that might have different movements.

DECLARATION OF CONFLICT OF INTEREST

Yusuke Satonaka, Ryosuke Wakamoto, and Michitaka Iida are employees of Information Services International-Dentsu, LTD, which supports digital transformation with solid technological and creative capabilities, and provided the programming code for the performance analysis in this study. To eliminate the possibility of bias, investigators were not involved in any data handling procedures.

FUNDING

This research was supported in part by grants from the Advanced Research Initiative for Human High Performance (ARIHHP), University of Tsukuba [grant number 2022(I)6], and by a JSPS KAKENHI Grant [grant number 22K17747].

REFERENCES

Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T., Zhang, F., & Grundmann, M. (2020). Blazepose: On-device real-time body pose tracking. *arXiv preprint arXiv:2006.10204.* https://doi.org/10.48550/arXiv.2006.10204

Brady, C., Tuyls, K., & Omidshafiei, S. (2021). *AI for Sports*. CRC Press.

Cai, J., Hu, J., Tang, X., Hung, T.-Y., & Tan, Y.-P. (2020). Deep historical long short-term memory network for action recognition. *Neurocomputing, 407*, 428-438. https://doi.org/10.1016/j.neucom.2020.03.111

Cust, E. E., Sweeting, A. J., Ball, K., & Robertson, S. (2019). Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance. *Journal of Sports Sciences, 37*(5), 568-600. https://doi.org/10.1080/02640414.2018.1521769

Devanandan, M., Rasaratnam, V., Anbalagan, M. K., Asokan, N., Panchendrarajan, R., & Tharmaseelan, J. (2021). Cricket Shot Image Classification Using Random Forest. In *2021 3rd International Conference on Advancements in Computing (ICAC)* (pp. 425-430). Colombo, Sri Lanka. https://doi.org/10.1109/ICAC54203.2021.9671109.

Kovalchik, S. A., & Reid, M. (2017). Comparing matchplay characteristics and physical demands of junior and professional tennis athletes in the era of big data. *Journal of sports science & medicine, 16*(4), 489-497. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5721178/

Ó Conaire, C., Connaghan, D., Kelly, P., O'Connor, N. E., Gaffney, M., & Buckley, J. (2010). Combining inertial and visual sensing for human action recognition in tennis. *Proceedings of the first ACM international workshop on Analysis and retrieval of tracked events and motion in imagery streams*, 51-56. https://doi.org/10.1145/1877868.1877882

O'Donoghue, P. (2005). Normative Profiles of Sports Performance. *International Journal of Performance Analysis in Sport, 5*(1), 104-119. https://doi.org/10.1080/24748668.2005.11868319

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 26*(1), 43-49. https://doi.org/10.1109/TASSP.1978.1163055

Shah, H., Chokalingam, P., Paluri, B., Pradeep, N., & Raman, B. (2007). Automated stroke classification in tennis. In *Image Analysis and Recognition: 4th International Conference, ICIAR 2007, Montreal, Canada, August 22-24, 2007. Proceedings* 4 (pp. 1128-1137). Springer Berlin Heidelberg.

Takahashi, H., Okamura, S., & Murakami, S. (2022). Performance analysis in tennis since 2000: A systematic review focused on the methods of data collection. *International Journal of Racket Sports Science, 4*(2), 40-55. https://doi.org/10.30827/Digibug.80900

Yoshida, S., Kitajima, E., & Miyata, R. (2021). Analyzing the motion of service on tennis using a pose-estimation model. In *IEICE Conferences Archives* [in Japanese].

Zhu, G., Xu, C., Huang, Q., & Gao, W. (2006). Action recognition in broadcast tennis video. *18th International Conference on Pattern Recognition (ICPR'06)*, 1, 251-254.