Electronic Research Archive

http://www.aimspress.com/journal/era

ERA, 34(10): 7739–7778.
DOI: 10.3934/era.2026333
Received: 08 June 2026
Revised: 21 July 2026
Accepted: 10 August 2026
Published: 08 September 2026

Research article

An explainable AI framework for badminton stroke quality assessment via hierarchical fuzzy inference and LLM-assisted coaching analytics

Hsiang-Chieh Chen[1,*], Hsin-Lun Liu[1], Jun-Lung Hsiao[2] and Pei-Yuan Hung[2]

[1] Department of Mechanical Engineering, National Central University, Taoyuan City, Taiwan
[2] Department of Electrical Engineering, National United University, Miaoli County, Taiwan

* **Correspondence:** Email: hcchen@cc.ncu.edu.tw.

Abstract: This study proposes an explainable computational intelligence framework for badminton stroke-quality assessment with large language model (LLM)-assisted coaching analytics. The proposed framework integrates stroke event segmentation, skeletal motion analysis, graph-based action recognition, stroke-quality assessment, and natural-language feedback generation. Racket–shuttlecock contact moments are first detected to extract individual stroke clips. MediaPipe Pose is employed to obtain skeletal joint trajectories, from which six stroke performance indices are formulated, including swing velocity index, wrist trajectory smoothness index, wrist posture index, elbow posture index, footwork efficiency index, and racket control index. These indices are integrated to generate interpretable stroke-quality ratings and radar-chart visualizations. For stroke classification, the proposed mST-GCN++ model achieved 94.34% accuracy across 11 badminton stroke categories, while the racket–shuttlecock contact detection module achieved an F1-score of 99.74%. The recognized stroke type, stroke performance indices, and assessment results are subsequently provided to an LLM-based module to generate personalized coaching feedback. Experimental results demonstrate that the proposed framework can effectively distinguish effective and ineffective stroke executions through interpretable quantitative assessments. Human-in-the-loop evaluation with badminton experts and players achieved an average score of 8.594/10, indicating that the generated feedback is technically accurate, consistent with observed movements, and useful for practical training. The proposed framework contributes by introducing six interpretable stroke performance indices that bridge low-level skeletal motion and high-level coaching knowledge, a hierarchical fuzzy inference system that performs transparent biomechanical reasoning for stroke-quality assessment, and a knowledge-guided LLM-assisted coaching framework that generates grounded natural-language coaching feedback from structured assessment results.

Keywords: badminton stroke assessment; fuzzy inference system; human pose estimation; large language model; skeleton-based action recognition; spatio-temporal graph convolutional networks

1. Introduction

Badminton is a rapid and technically complex racket sport requiring coordinated upper-limb movement, racket control, footwork, and precise shot timing [1–3]. Stroke quality depends not only on successfully striking the shuttlecock but also on how the stroke is executed [1,2,4]. Swing velocity, wrist trajectory, joint motion, lower-body positioning, and racket-face orientation significantly influence shot accuracy, stability, and consistency [3,5]. Conventional badminton training is predominantly evaluated by coaches through visual assessment and experience. Although expert coaching is indispensable, consistent and personalized feedback is often unavailable to beginners, amateur players, and developmental teams [6,7]. Therefore, an accessible, objective, and interpretable assessment framework is needed to evaluate badminton strokes in practical training [8].

Recent advances in computer vision and deep learning facilitate sports video analysis through object detection, trajectory tracking, pose estimation, and action recognition [9,10]. In badminton, vision-based approaches have been applied to shuttlecock and racket detection, hit-event extraction, player movement analysis, and stroke classification [11–13]. However, badminton video analysis remains challenging due to the shuttlecock's small size and high speed, rapid racket motion, and abrupt changes in player posture. Deep learning detection [11,14,15] and TrackNet trajectory-tracking models [16,17] have demonstrated strong performance in shuttlecock localization and badminton event analysis, highlighting the extraction of stroke-related information, racket-shuttlecock contact moments, and stroke clips from video sequences.

Human pose estimation is key to vision-based badminton analysis [18–21]. It transforms video frames into skeletal joint coordinates, offering a compact representation of player movement [19,21]. Markerless frameworks, including OpenPose [22], MediaPipe Pose [23], BlazePose [24], and YOLO-Pose variants [18,25], are widely adopted for motion analysis as they eliminate the need for wearable sensors and laboratory-based motion-capture systems. In badminton, pose-based methods have analyzed joint angles, movement trajectories, distance patterns, and stroke-related body postures. Compared with stereo-vision and multi-camera systems, monocular and smartphone-based solutions offer lower hardware complexity and greater accessibility for routine training [21,26]. However, pose-estimation outputs alone lack direct interpretability for coaching and must be transformed into meaningful motion indicators for stroke-quality assessment.

Skeleton-based action recognition classifies badminton strokes from temporal joint-coordinate sequences [27–30]. Spatial-temporal graph convolutional networks (ST-GCNs) model body joints as graph nodes and anatomical-temporal connections as graph edges, learning spatial joint relationships and temporal motion dynamics [27]. Subsequent models, including adaptive graph convolutional networks [30,31], multi-stream attention-enhanced graph models [28], channel-wise topology refinement graph convolution [32], and temporal Transformer architectures [13,33,34], have improved skeleton-based action recognition by enhancing graph topology learning and temporal dependency modeling. In badminton and other racket sports, skeleton-based and Transformer-based architectures have enabled fine-grained stroke recognition, demonstrating the effectiveness of skeletal motion sequences in distinguishing badminton strokes [35,36].

Despite these advances, stroke recognition alone remains inadequate for coaching-oriented evaluation. Although recognition models classify strokes (e.g., 11 stroke categories, including forehand clear, forehand cut, forehand drive, forehand drop, forehand lift, backhand clear, backhand cut, backhand drive, backhand drop, backhand lift, and smash), they do not assess execution quality or technical flaws [37,38]. Practical training requires interpretable motion quality information, including wrist-trajectory stability, elbow positioning appropriateness near impact, footwork efficiency, and racket-face alignment [1,38–40]. Consequently, stroke recognition must integrate with interpretable motion indicators and quantitative scoring.

Numerical scores alone are also difficult for general users to interpret. Large language model (LLM)-based coaching studies show that LLMs can facilitate personalized natural-language feedback generation, but require safety, grounding, and reliability in health- and sports-related applications [41–44]. In sports training, LLMs should use measurable evidence rather than raw video to infer stroke quality [45,46]. Instead, feedback generation should be based on structured inputs, including recognized stroke type, motion indicators, radar-chart ratings, and observed movement deviations. Human-in-the-loop verification remains essential for technically meaningful AI-generated feedback suitable for practical coaching scenarios [46–48].

Recent advances in artificial intelligence (AI) have shifted from task-oriented prediction models to explainable, trustworthy, and human-centered intelligent systems. Emerging studies emphasize structured reasoning, information fusion, and human–AI collaboration to improve the transparency and reliability of AI-assisted decision support [49–51]. However, existing AI-based badminton analysis remains largely task-specific, focusing on stroke recognition, pose estimation, or motion analysis. Furthermore, natural-language feedback is typically generated from numerical features or predefined templates, without explicitly incorporating structured biomechanical reasoning and domain-specific coaching knowledge. Consequently, an explainable coaching framework that integrates vision-based inputs, transparent quantitative reasoning, and knowledge-guided coaching analytics remains largely unexplored.

To address this gap, this study proposes an explainable AI framework for badminton stroke-quality assessment that bridges vision-based motion analysis, computational intelligence, and LLM-assisted coaching analytics. The proposed framework explicitly transforms human motion into interpretable biomechanical representations through six stroke performance indices (SPIs). These SPIs are subsequently integrated by a hierarchical fuzzy inference system (HFIS) to provide explainable stroke-quality assessment, enabling quantitative reasoning from low-level motion features to high-level coaching interpretations. Finally, rather than directly prompting an LLM with raw video or unstructured descriptions, the proposed framework adopts a knowledge-guided coaching strategy in which the recognized stroke type, SPI representations, fuzzy assessment results, and radar-chart ratings serve as structured evidence for generating grounded natural-language feedback. A smartphone-based deployment is further implemented to demonstrate the practicality of the proposed framework in real-world badminton training scenarios. The main contributions of this study are summarized as follows:

• Six SPIs are introduced to explicitly represent badminton stroke execution from complementary biomechanical perspectives. These representations establish an explainable bridge between low-level skeletal motion features and high-level coaching knowledge.

• An HFIS is developed to hierarchically aggregate the six SPIs into intermediate assessments and an overall stroke-quality score, providing transparent and interpretable reasoning rather than direct score regression.

• A knowledge-guided LLM-assisted coaching module is presented to generate grounded and personalized natural-language feedback from structured assessment results, integrating domain-specific badminton knowledge with explainable AI outputs.

• Overall, an explainable badminton stroke-quality assessment framework is proposed, and its effectiveness and practical applicability are evaluated through experiments, including stroke recognition, statistical agreement analysis, binary classification of effective and ineffective strokes, and human-in-the-loop evaluation.

The remainder of this paper is organized as follows: Section 2 presents the overall system architecture, deployment pipeline, and human-in-the-loop design of the proposed framework. Section 3 details the vision-based motion analysis framework, covering stroke event segmentation, pose estimation, effective-hitting-interval localization, biomechanical motion feature extraction, and stroke classification. Section 4 presents the stroke-quality assessment framework, including SPI construction, hierarchical fuzzy inference, and LLM-assisted coaching analytics. Section 5 reports the experimental setup, quantitative results, representative assessment cases, statistical agreement analysis, effective and ineffective stroke classification, and human-in-the-loop evaluation outcomes. Finally, Section 6 concludes the paper and discusses future research directions.

2. System overview

2.1. Problem setting and system definition

This study aims to develop a practical, coach-aligned framework for badminton stroke-quality assessment using vision-based motion analysis and widely available devices such as smartphones. Unlike conventional sports vision systems that primarily emphasize action recognition, the proposed approach focuses on quantitative performance assessment and explainable coaching support. To characterize stroke execution, six SPIs are formulated and then integrated through a hierarchical fuzzy inference system to generate interpretable stroke-quality assessments. Rather than replacing human coaches, the proposed framework serves as an explainable decision-support tool that assists coaches and players by providing quantitative performance analytics and personalized training insights.

2.2. Overall framework

Figure 1 illustrates the overall architecture of the proposed badminton stroke-quality assessment framework. Unlike conventional action-recognition pipelines, the system converts continuous badminton training videos into stroke-level evaluations and coaching feedback through seven functional modules, as described below.

1) Stroke event segmentation: Badminton videos are recorded using a smartphone or single-camera setup to ensure practical usability. As continuous recordings may include multiple strokes and non-hitting intervals, stroke segmentation is first performed to isolate individual stroke clips. The system tracks shuttlecock and racket positions across frames and estimates the impact moments based on their spatial proximity. A short temporal window centered on the detected contact frame is then extracted as an individual stroke clip.

2) Human pose estimation: Each extracted stroke clip is processed using pose estimation to obtain skeletal body joint key points frame-by-frame, producing temporal joint-coordinate sequences that

describe player motion for subsequent motion analysis.
3) Stroke motion representation: The skeletal sequences are transformed into biomechanical motion features, including swing velocity, wrist trajectory, wrist posture angle, elbow posture angle, footwork patterns, and racket-face characteristics. These features provide quantitative representations of stroke execution and serve as the basis for subsequent SPI construction.
4) Stroke type recognition: The system classifies each clip into one of 11 stroke categories using skeletal motion features. These categories include forehand and backhand clear, cut, drive, drop, lift, and smash. The predicted stroke type provides context for stroke-specific assessment.
5) Hitting interval localization: The racket–shuttlecock contact moment is detected from their spatial proximity. A temporal window centered on the detected contact frame extracts the stroke clip, and the effective hitting interval is further localized using wrist motion. This interval provides the temporal basis for motion feature extraction and stroke-quality scoring.
6) Stroke-quality assessment: The extracted biomechanical motion features are converted into six SPIs. Then, a hierarchical fuzzy inference system integrates the SPIs into intermediate biomechanical assessments and an overall stroke-quality score, providing interpretable and explainable evaluation results.
7) LLM-assisted analytics: The recognized stroke type, six SPIs, fuzzy assessment results, and radar-chart ratings are supplied to the LLM-assisted coaching analytics module to generate personalized natural-language coaching feedback. The feedback explains stroke performance in an interpretable manner through an interactive interface, helping players and coaches review strengths, weaknesses, and potential improvements.

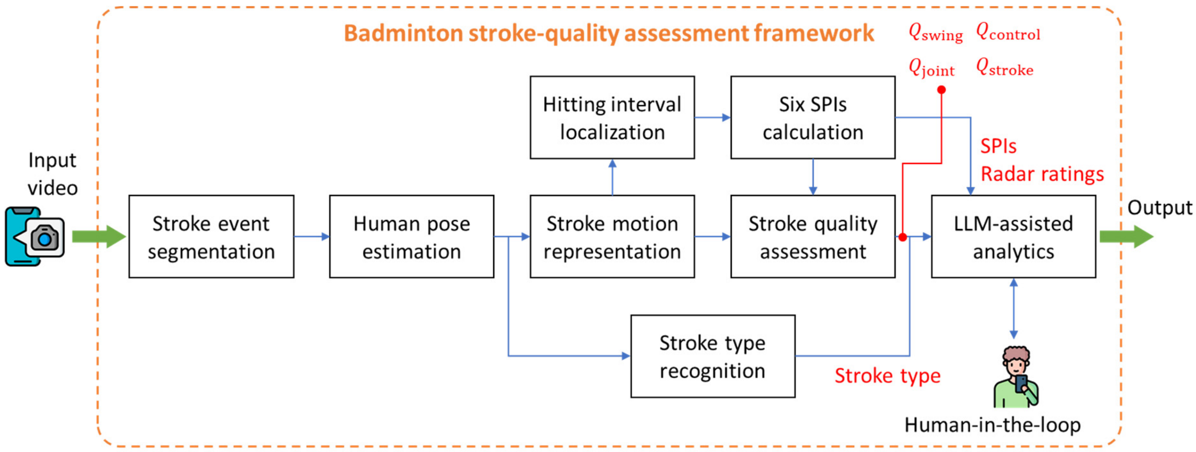


Figure 1. The proposed badminton stroke-quality assessment framework.

2.3. Deployment and human-in-the-loop design

The proposed framework is deployed through a smartphone application that provides a convenient interface for badminton training. Users first record their badminton strokes and may optionally trim the video to retain the relevant stroke segment before uploading it to the cloud server. The application is responsible only for video acquisition, trimming, and uploading; it does not perform any motion analysis or allow users to manually specify analysis parameters. After the video is uploaded, the complete analysis pipeline described in Subsection 2.2 is executed automatically on the cloud server. The resulting radar-chart scores and personalized coaching feedback are then returned to the smartphone interface, as illustrated in Figure 2. Players and coaches can then review the quantitative assessment results, identify technical strengths and weaknesses, and obtain interpretable training recommendations.

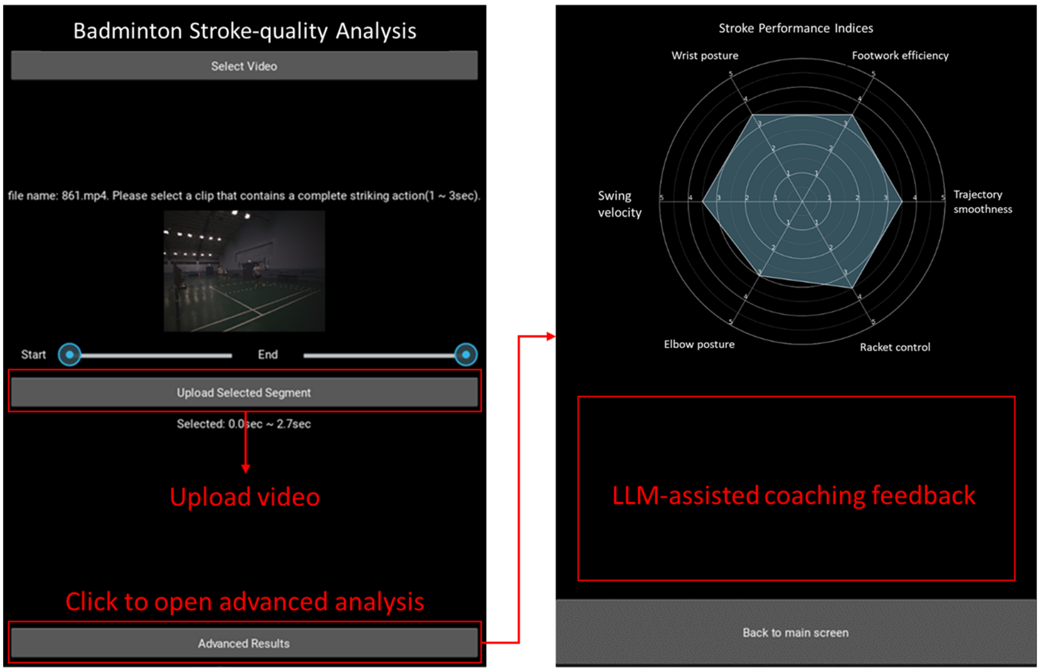


Figure 2. Smartphone interface of the proposed badminton stroke-quality assessment.

3. Vision-based stroke analysis

This section presents the motion analysis framework for transforming raw badminton stroke videos into structured motion representations. The representation comprises stroke event segmentation, human pose estimation, effective hitting interval localization, biomechanical motion feature extraction, and stroke type recognition.

3.1. Stroke event segmentation

To segment continuous badminton videos into individual stroke clips, a vision-based event segmentation method is employed to detect the racket–shuttlecock contact moments using YOLO-based object detection. Let $c_{\mathrm{r},t} = (u_{\mathrm{r},t}, v_{\mathrm{r},t})$ and $c_{\mathrm{s},t} = (u_{\mathrm{s},t}, v_{\mathrm{s},t})$ denote the center coordinates of the racket and shuttlecock within frame t, respectively. The coordinates are normalized by the image width and height before distance computation. The spatial distance between the racket and the shuttlecock is computed as

$$d_{\mathrm{rs},t} = \sqrt{(u_{\mathrm{r},t} - u_{\mathrm{s},t})^2 + (v_{\mathrm{r},t} - v_{\mathrm{s},t})^2} \tag{1}$$

The contact moment generally corresponds to the frame where the racket and shuttlecock exhibit minimum spatial separation. Accordingly, this moment is determined as

$$t_{\text{contact}} = \arg\min_{t\in\mathcal{T}_{\text{candi}}} d_{\text{rs},t}\,,\ \mathcal{T}_{\text{candi}} = \{t \mid d_{\text{rs},t} \le d_{\text{Th}}\} \tag{2}$$

where $\mathcal{T}_{\text{candi}}$ is the set of valid contact-frame candidates, and $d_{\text{Th}} = 0.1$ is a predefined distance threshold. If no frame satisfies the threshold condition, the segment is rejected as an invalid contact event.

Once t_{contact} is identified, a temporal window centered on the frame is extracted as a stroke clip. The clip spans $[t_{\text{contact}} - T_{\text{pre}}, t_{\text{contact}} + T_{\text{post}}]$, where T_{pre} and T_{post} denote the numbers of frames before and after the contact frame, respectively. The resulting clip length is $L_{\text{clip}} = T_{\text{pre}} + T_{\text{post}} + 1$. In this study, $T_{\text{pre}} = 20$ and $T_{\text{post}} = 20$ are empirically selected for videos recorded at 30 frames per second to cover the preparation, swing, and follow-through phases. Figure 3 illustrates the YOLO-based racket and shuttlecock detection result used for stroke event segmentation, where detected object centers are employed to estimate the racket–shuttlecock contact moment.

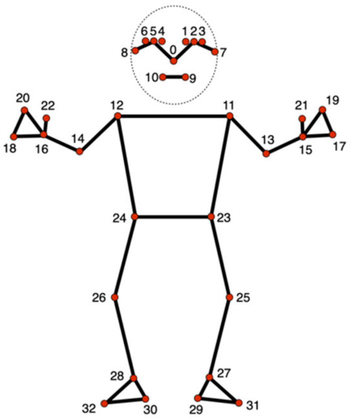
Figure 3. Result of YOLO-based racket and shuttlecock detection.

3.2. Human pose estimation

Each segmented stroke clip is processed using MediaPipe Pose [23] to detect 33 body key points (indexed 0–32) per frame, providing 2D image coordinates and approximate relative depth from a monocular camera. These key points (Figure 4) form the skeletal representation used to derive the motion features (Subsection 3.4).

Given a stroke clip of N frames, MediaPipe Pose extracts 33 body key points per frame t, denoted as $P_t = \{\mathbf{p}_{0,t}, \mathbf{p}_{1,t}, \ldots, \mathbf{p}_{32,t}\}$, where each key point $\mathbf{p}_{i,t} = (x_{i,t}, y_{i,t}, z_{i,t})$ represents the coordinates of the i-th body joint, with $i = 0,1,\ldots,32$. The x-axis corresponds to the horizontal image direction, the y-axis corresponds to the vertical image direction, and the z-axis encodes the estimated depth direction relative to the camera.

Pose estimation across all frames yields a temporal skeletal sequence $S = \{P_1, P_2, \ldots, P_N\}$. An extended Kalman filter [52] applied to the extracted joint trajectories enhances temporal consistency. Body-centered coordinate normalization (subtracting the hip midpoint) and shoulder-width-based scale normalization (dividing by the shoulder width, i.e., the distance between the left and right shoulder joints) are employed to reduce variations due to camera distance and body size differences.



0	nose	11	left shoulder	22	right thumb
1	left eye (inner)	12	right shoulder	23	left hip
2	left eye	13	left elbow	24	right hip
3	left eye (outer)	14	right elbow	25	left knee
4	right eye (inner)	15	left wrist	26	right knee
5	right eye	16	right wrist	27	left ankle
6	right eye (outer)	17	left pinky	28	right ankle
7	left ear	18	right pinky	29	left heel
8	right ear	19	left index	30	right heel
9	mouth (left)	20	right index	31	left foot index
10	mouth (right)	21	left thumb	32	right foot index

Figure 4. Human skeletal key points detected using MediaPipe Pose [23].

3.3. Effective hitting interval localization

Although stroke event segmentation extracts individual stroke clips, each clip may still include frames outside the active swing motion. Therefore, the effective hitting interval is identified to localize the main swing phase using frame-by-frame wrist-joint displacement of the hitting hand. Let $\mathbf{p}_{\text{wri},t}$ denote the wrist coordinate at frame t (key point 15 for left-handed players and key point 16 for right-handed players, as defined in Figure 4). The per-frame wrist displacement is computed as

$$d_{\text{wri},t} = \left\|\mathbf{p}_{\text{wri},t} - \mathbf{p}_{\text{wri},t-1}\right\|_2 \tag{3}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. For a clip containing N frames, the mean and standard deviation of the wrist displacement are calculated as

$$\mu = \frac{1}{N-1}\sum_{t=2}^{N} d_{\text{wri},t} \ \text{ and } \ \sigma = \sqrt{\frac{1}{N-1}\sum_{t=2}^{N}\left(d_{\text{wri},t} - \mu\right)^2} \tag{4}$$

The candidate swing frames are then empirically defined as $\mathcal{T}_{\text{swing}} = \{\, t \mid d_{\text{wri},t} \geq \mu + 3\sigma \,\}$. The threshold $\mu + 3\sigma$ was empirically determined through preliminary experiments and was found to effectively isolate the active swing phase while suppressing minor wrist fluctuations. The effective hitting interval $[t_{\text{sta}}, t_{\text{end}}]$ is defined as the continuous segment in $\mathcal{T}_{\text{swing}}$ that contains the detected contact moment t_{contact}. This constraint ensures consistency between the effective hitting interval and the contact frame detected in Subsection 3.1. Here, t_{sta} and t_{end} denote the start and end frames of the active swing motion, respectively. If no candidate segment contains t_{contact}, the closest candidate segment to t_{contact} is selected, or the clip is rejected as an invalid swing sample.

3.4. Stroke motion representation

To characterize stroke execution, six biomechanical motion features are first introduced from skeletal joint trajectories within the effective hitting interval to capture swing dynamics, upper-limb kinematics, footwork, racket orientation, and body movement. These features provide interpretable kinematic descriptors for subsequent SPI construction in Subsection 4.1.

1) Swing velocity: Swing velocity is estimated from wrist velocity within the effective hitting interval. Peak wrist velocity represents swing intensity, as it reflects the maximum speed of the racket arm during stroke execution. It is defined as

$$\hat{V}_{\text{swing}} = \max_{t \in [t_{\text{sta}}, t_{\text{end}}]} V_t \tag{5}$$

where $V_t = d_{\text{wri},t}/\Delta t$ denotes the wrist velocity at frame t, $d_{\text{wri},t}$ is the wrist displacement defined in Eq (3), and Δt represents the time interval between consecutive frames.
2) Wrist trajectory: The wrist trajectory represents the spatial path of the hitting arm during the effective hitting interval. Given the wrist coordinate $\mathbf{p}_{\text{wri},t}$, the wrist trajectory is defined as $\{\mathbf{p}_{\text{wri},t} \mid \forall\, t_{\text{sta}} \le t \le t_{\text{end}}\}$. Its geometric pattern describes the swing path of the racket arm and is utilized to evaluate motion consistency and stroke control.
3) Wrist joint angle: The wrist joint angle describes wrist-related articulation of the hitting hand during stroke execution. Let $\mathbf{p}_{\text{elb},t}$ denote the elbow key point, $\mathbf{p}_{\text{wri},t}$ the wrist key point, and $\mathbf{p}_{\text{hand},t}$ the midpoint of index and pinky key points at frame t. The wrist angle is determined as

$$\theta_{\text{wri},t} = \cos^{-1}\left(\frac{(\mathbf{p}_{\text{elb},t}-\mathbf{p}_{\text{wri},t})\cdot(\mathbf{p}_{\text{hand},t}-\mathbf{p}_{\text{wri},t})}{\|\mathbf{p}_{\text{elb},t}-\mathbf{p}_{\text{wri},t}\|\|\mathbf{p}_{\text{hand},t}-\mathbf{p}_{\text{wri},t}\|}\right) \tag{6}$$

This angle approximates wrist control and wrist snap during the stroke, with left or right key points selected according to the player's hitting hand.
4) Elbow joint angle: The elbow joint angle describes the upper arm–forearm coordination of the hitting arm. Let $\mathbf{p}_{\text{sho},t}$, $\mathbf{p}_{\text{elb},t}$, and $\mathbf{p}_{\text{wri},t}$ denote the shoulder, elbow, and wrist joint coordinates, respectively. The elbow joint angle is computed as

$$\theta_{\text{elb},t} = \cos^{-1}\left(\frac{(\mathbf{p}_{\text{sho},t}-\mathbf{p}_{\text{elb},t})\cdot(\mathbf{p}_{\text{wri},t}-\mathbf{p}_{\text{elb},t})}{\|\mathbf{p}_{\text{sho},t}-\mathbf{p}_{\text{elb},t}\|\|\mathbf{p}_{\text{wri},t}-\mathbf{p}_{\text{elb},t}\|}\right) \tag{7}$$

This angle feature reflects arm extension, flexion, and upper-limb coordination during stroke execution.

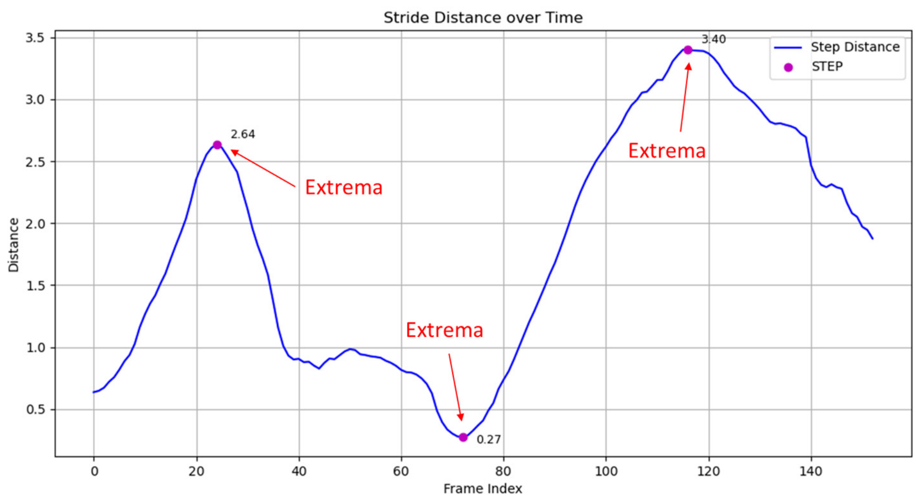


Figure 5. Example of step-count estimation using the foot-distance variation curve.

5) Foot trajectory and step pattern analysis: Foot movement is analyzed to describe lower-body positioning and movement efficiency during the stroke. Let the left and right ankle key points (extracted from human pose estimation) at frame t represent foot positions $\mathbf{p}_{\mathrm{Lf},t}$ and $\mathbf{p}_{\mathrm{Rf},t}$, respectively, on an image-based movement plane. The distance between the two feet at frame t is defined as

$$d_{\mathrm{foot},t} = \left\|\mathbf{p}_{\mathrm{Lf},t} - \mathbf{p}_{\mathrm{Rf},t}\right\|_2 \tag{8}$$

Tracking $d_{\mathrm{foot},t}$ within the effective hitting interval yields a foot-distance variation curve (Figure 5). Local extrema correspond to large strides or cross-steps, and local minima indicate moments when feet move closer together. The number of detected extrema within the hitting interval estimates the step count and characterizes the player's footwork pattern during stroke execution.

6) Racket face analysis: Racket orientation at the contact moment is crucial for shuttlecock direction and stroke control. Since MediaPipe Pose does not provide racket key points, this study approximates racket-face orientation using hand key points. Let $\mathbf{p}_{\mathrm{wri},t}$, $\mathbf{p}_{\mathrm{pin},t}$, and $\mathbf{p}_{\mathrm{ind},t}$ denote the wrist, pinky, and index finger key points of the hitting hand, respectively (key points 16, 18, and 20 for right-handed players; 15, 17, and 19 for left-handed players). The midpoint of the pinky and index-finger key points is computed as

$$\mathbf{p}_{\mathrm{mid},t} = \frac{\mathbf{p}_{\mathrm{pin},t}+\mathbf{p}_{\mathrm{ind},t}}{2} \tag{9}$$

The approximate hand–racket shaft direction is then defined as

$$\mathbf{u}_{\mathrm{rs},t} = \frac{\mathbf{p}_{\mathrm{mid},t}-\mathbf{p}_{\mathrm{wri},t}}{\left\|\mathbf{p}_{\mathrm{mid},t}-\mathbf{p}_{\mathrm{wri},t}\right\|} \tag{10}$$

An approximate racket-center point is estimated by extending this direction from the wrist:

$$\mathbf{p}_{\mathrm{rc},t} = \mathbf{p}_{\mathrm{wri},t} + l_{\mathrm{est}}\,\mathbf{u}_{\mathrm{rs},t} \tag{11}$$

where l_{est} denotes the estimated racket-shaft length obtained through calibration. The hand-plane normal is estimated using the wrist, index, and pinky key points:

$$\mathbf{n}_{\mathrm{hand},t} = \frac{(\mathbf{p}_{\mathrm{ind},t}-\mathbf{p}_{\mathrm{wri},t})\times(\mathbf{p}_{\mathrm{pin},t}-\mathbf{p}_{\mathrm{wri},t})}{\left\|(\mathbf{p}_{\mathrm{ind},t}-\mathbf{p}_{\mathrm{wri},t})\times(\mathbf{p}_{\mathrm{pin},t}-\mathbf{p}_{\mathrm{wri},t})\right\|} \tag{12}$$

The lateral direction of the approximated racket face is defined as

$$\mathbf{u}_{\mathrm{rf},t} = \frac{\mathbf{u}_{\mathrm{rs},t}\times\mathbf{n}_{\mathrm{hand},t}}{\left\|\mathbf{u}_{\mathrm{rs},t}\times\mathbf{n}_{\mathrm{hand},t}\right\|} \tag{13}$$

Finally, the approximate racket-face normal is estimated as

$$\mathbf{n}_{\mathrm{rac},t} = \frac{\mathbf{u}_{\mathrm{rs},t}\times\mathbf{u}_{\mathrm{rf},t}}{\left\|\mathbf{u}_{\mathrm{rs},t}\times\mathbf{u}_{\mathrm{rf},t}\right\|} \tag{14}$$

This vector serves as a racket-face orientation proxy for evaluating the consistency between the estimated hand–racket orientation and the reference stroke pattern at the contact frame. Figure 6 illustrates this geometric construction for a right-handed player using key points 16 (wrist), 18 (pinky), and 20 (index finger) to approximate hand–racket shaft direction, hand-plane normal, and racket-face orientation. For left-handed players, key points 15, 17, and 19 are used instead.

It should be noted that the proposed framework is designed for smartphone-based deployment, where the recorded videos primarily capture the player's full-body motion and the overall court scene. Consequently, the proposed method does not aim to recover the absolute racket-face pose with high geometric accuracy. Instead, the estimated racket orientation is used as a reference for evaluating the temporal consistency of racket motion throughout the stroke. Therefore, the proposed formulation emphasizes the stability of racket orientation rather than its absolute geometric accuracy. Although whole-body pose estimation may introduce localization errors for hand landmarks, these errors have relatively limited influence on the proposed stability-based assessment, as the evaluation primarily relies on relative temporal variation rather than on instantaneous orientation estimates.

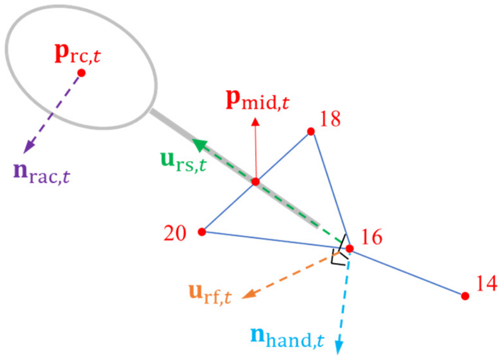


Figure 6. Geometric construction of the racket-face orientation proxy for a right-handed player.

3.5. Stroke type recognition

Stroke type recognition identifies the stroke category of each segmented clip to support subsequent quality assessment. Using the skeletal motion sequence $S = \{P_1, P_2, \dots, P_N\}$ as input with N frames, with 33 key points $P_t = \{\mathbf{p}_{0,t}, \mathbf{p}_{1,t}, \dots, \mathbf{p}_{32,t}\}$ per frame from Subsection 3.2, a spatio-temporal graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is constructed, where $\mathcal{V}$ denotes body-joint nodes, and $\mathcal{E}$ represents graph edges. Spatial edges encode anatomical connections between joints within a frame, and temporal edges connect the same joints across successive frames to describe motion dynamics. Based on this graph representation, stroke recognition is performed using the modified ST-GCN++ (mST-GCN++) model, which integrates an MS-TCN–Transformer temporal module into the ST-GCN++ backbone [53]. Eleven stroke categories are defined: forehand clear, forehand cut, forehand drive, forehand drop, forehand lift, backhand clear, backhand cut, backhand drive, backhand drop, backhand lift, and smash. The recognized stroke type subsequently enables stroke-specific motion interpretation and expert-aligned quality assessment.

3.5.1. The architecture of the proposed mST-GCN++ model

ST-GCN++ is a variant of ST-GCN, developed in [53], with modifications to the architectures of the spatial and temporal modules. The proposed mST-GCN++ model builds upon the ST-GCN++ backbone for skeleton-based stroke recognition. In this framework, badminton stroke clips are modeled as spatio-temporal graphs, in which body joints serve as nodes, and anatomical and temporal connections define the edges. This representation enables the model to capture both spatial dependencies among joints and temporal motion dynamics within skeletal sequences.

The baseline ST-GCN++ model incorporates a multi-branch MS-TCN temporal module [54] (Figure 7). The input features are first processed using a 1×1 convolution and then fed into multiple temporal branches, including dilated temporal convolution, max-pooling, and standard convolution. The outputs from all branches are concatenated and fused to form the final temporal representation. This design enables the model to capture local motion patterns across multiple temporal receptive fields.

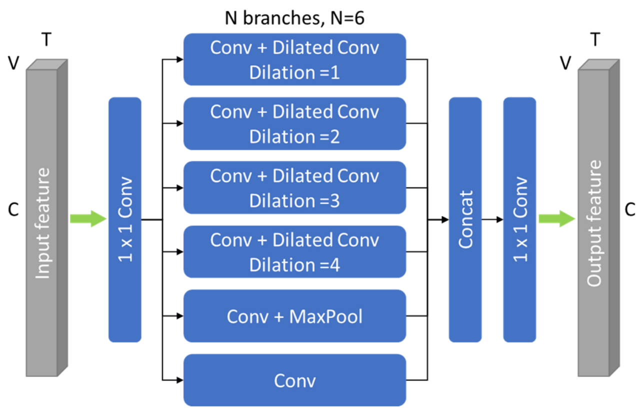


Figure 7. Architecture of the original MS-TCN temporal module in ST-GCN++.

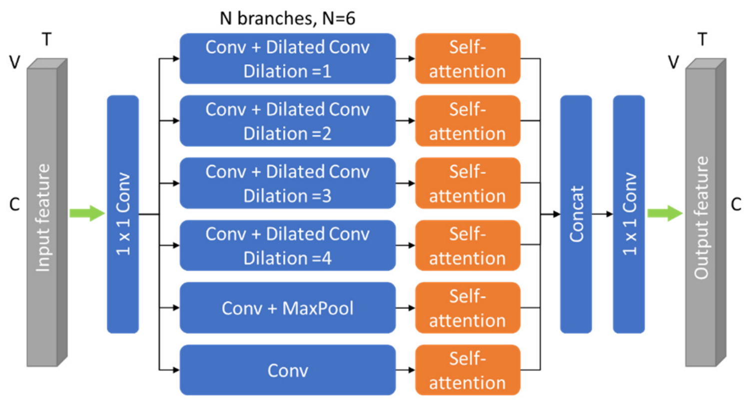


Figure 8. Architecture of the modified MS-TCN module with branch-wise attention.

To improve temporal feature discrimination, a modified MS-TCN variant with branch-wise attention is also explored, as shown in Figure 8. The attention mechanism enhances informative temporal responses while suppressing less relevant features, particularly around key stroke phases such as racket-arm acceleration and impact. This variant retains the original multi-branch MS-TCN structure while improving its ability to focus on discriminative stroke-related motion patterns.

In addition, PatchTST is compared with global temporal modeling [55]. Unlike MS-TCN-based modules, PatchTST segments the skeletal sequence into temporal patches and employs a Transformer encoder to model dependencies across temporal segments. However, lacking the original MS-TCN branch structure, it serves as a comparison baseline rather than the final architecture.

The final model incorporates an MS-TCN–Transformer temporal module, as illustrated in Figure 9.

In this design, the MS-TCN component first extracts local and multi-scale temporal features, which are then fed into a Transformer encoder to capture long-range temporal dependencies. This sequential architecture enables the model to jointly model local stroke dynamics and global temporal relationships within the motion sequence.

Overall, the original MS-TCN serves as the baseline temporal module; the attention-enhanced MS-TCN improves local temporal discrimination; and PatchTST offers a Transformer-based temporal comparison. The final MS-TCN–Transformer module, integrating multi-scale temporal convolution with global dependency modeling, is adopted as the temporal component of the proposed mST-GCN++ for badminton stroke type recognition.

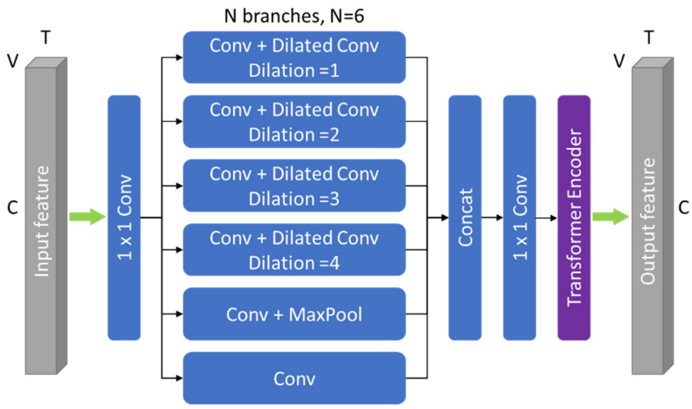


Figure 9. Architecture of the proposed MS-TCN–Transformer temporal module.

4. Computational intelligence–based stroke-quality assessment

The proposed framework evaluates badminton stroke quality through a multi-stage assessment that integrates biomechanical motion analysis, interpretable performance representations, and LLM-assisted coaching analytics. Motion features, extracted from skeletal joint trajectories, are first transformed into six SPIs that capture complementary aspects of stroke execution, including swing dynamics, upper-limb posture, footwork efficiency, and racket control. These SPIs are then integrated using a hierarchical fuzzy inference model to produce interpretable stroke-quality ratings. The resulting assessment outcomes are used to generate radar-chart visualizations and personalized coaching feedback via an LLM, enabling interpretable and actionable training recommendations.

4.1. Stroke performance index (SPI) construction

The six motion features introduced in Subsection 3.4 characterize stroke execution from multiple biomechanical perspectives, including swing dynamics, upper-limb kinematics, racket control, and footwork. Each feature represents a specific functional aspect of badminton stroke performance and provides an interpretable measure of movement quality. To facilitate quantitative assessment, the extracted motion features are transformed into a unified stroke performance index (SPI) representation. Specifically, each feature is converted into a corresponding SPI, denoted by $s_k \in [0,1]$ for $k = 1,2, \ldots ,6$, where higher values indicate better performance. The resulting SPIs are subsequently mapped onto a five-point rating scale for visualization and coaching feedback generation.

• SPI-1: Swing velocity index. This index evaluates racket-side wrist velocity appropriateness during the effective hitting interval. Since both excessively slow or excessively fast swings reduce stroke consistency, this index is computed from the deviation between the observed peak wrist velocity and a reference velocity:

$$s_1 = \exp\left(-\frac{\left(\hat{V}_{\text{swing}} - V_{\text{ref}}\right)^2}{2\sigma_V^2}\right) \tag{15}$$

where $\hat{V}_{\text{swing}}$ denotes the maximum wrist velocity within the effective hitting interval, as defined in Eq (5), V_{ref} represents the reference peak wrist velocity of a well-executed stroke, and σ_V controls velocity deviation tolerance. In practice, V_{ref} and σ_V can be derived from expert demonstrations, coach criteria, or the player's historical well-executed strokes. In this study, the reference values were obtained from expert demonstrations: $V_{\text{ref}} = 15\ \text{m/s}$ and $\sigma_V = 1.05$. This index is bounded within $[0,1]$, with higher values indicating swing velocity closer to the reference pattern. Stroke quality is not solely determined by maximum swing velocity. Excessively high swing velocities may reduce stroke stability and consistency. Therefore, the proposed scoring function favors swing velocities that are close to the reference pattern rather than simply rewarding higher velocities.

• SPI-2: Wrist trajectory smoothness index. This index quantifies the smoothness of racket-side wrist motion during the effective hitting interval. Let $\mathbf{p}_{\text{wri},t}$ denote the wrist position at frame t, and let N be the number of frames in the interval. The trajectory smoothness error is computed as follows.

$$e_{\text{traj}} = \frac{1}{N-2}\sum_{t=2}^{N-1}\left\|\mathbf{p}_{\text{wri},t+1} - 2\mathbf{p}_{\text{wri},t} + \mathbf{p}_{\text{wri},t-1}\right\|_2 \tag{16}$$

A smaller e_{traj} indicates a smoother and more stable wrist trajectory. The corresponding SPI-2 is defined as

$$s_2 = \exp\left(-\frac{e_{\text{traj}}^2}{2\sigma_{\text{traj}}^2}\right) \tag{17}$$

where σ_{traj} controls scoring sensitivity, set to $\sigma_{\text{traj}} = 0.3\max\left(e_{\text{traj}}^{\text{hist}}\right)$ determined from the player's historical data, where $e_{\text{traj}}^{\text{hist}}$ denotes the trajectory smoothness errors computed from the player's past strokes.

• SPI-3: Wrist posture index. This index assesses the consistency of wrist posture near impact relative to the reference movement pattern. Let $\theta_{\text{wri},t_{\text{contact}}}$ denote the wrist joint angle at the contact frame, and let $\theta_{\text{wri},t_{\text{contact}}}^{\text{ref}}$ refer to the corresponding reference angle. The SPI-3 is defined as

$$s_3 = \exp\left(-\frac{\left(\theta_{\text{wri},t_{\text{contact}}} - \theta_{\text{wri},t_{\text{contact}}}^{\text{ref}}\right)^2}{2\sigma_{\text{wri}}^2}\right) \tag{18}$$

where σ_{wri} controls the tolerance to wrist angle deviation. The reference angle and σ_{wri} can be determined from expert demonstrations, coach-defined criteria, or the player's historical well-executed strokes. In our work, $\sigma_{\text{wri}} = 1.2$ and $\theta_{\text{wri},t_{\text{contact}}}^{\text{ref}} = 165°$.

• SPI-4: Elbow posture index. This index evaluates whether arm posture near impact is consistent with the reference movement pattern. Let $\theta_{\text{elb},t_{\text{contact}}}$ denote the elbow joint angle at the contact frame, and let $\theta_{\text{elb},t_{\text{contact}}}^{\text{ref}}$ represent the corresponding reference angle. The normalized score is defined as

$$s_4 = \exp\left(-\frac{\left(\theta_{\mathrm{elb},t_{\mathrm{contact}}}-\theta_{\mathrm{elb},t_{\mathrm{contact}}}^{\mathrm{ref}}\right)^2}{2\sigma_{\mathrm{elb}}^2}\right) \tag{19}$$

where σ_{elb} controls the tolerance to elbow angle deviation. The reference angle and σ_{elb} can be derived from expert demonstrations, coach-defined criteria, or the player's historical well-executed strokes. In our work, $\sigma_{\mathrm{elb}} = 1.2$ and $\theta_{\mathrm{elb},t_{\mathrm{contact}}}^{\mathrm{ref}} = 172°$. This SPI ranges from 0 to 1, with higher values indicating an elbow posture closer to the reference at impact.

• SPI-5: Footwork efficiency index. This index evaluates lower-body movement during the effective hitting interval by considering both step pattern and movement efficiency. Let n_{step} denote the number of detected step-related events, and let $n_{\mathrm{step,ref}}$ represent the reference step count for an efficient stroke. The step-based score is defined as

$$s_{5,\mathrm{step}} = \exp\left(-\frac{\left(n_{\mathrm{step}}-n_{\mathrm{step,ref}}\right)^2}{2\sigma_{\mathrm{step}}^2}\right) \tag{20}$$

where $\sigma_{\mathrm{step}} = 1$ controls the tolerance to step-count variation. A higher $s_{5,\mathrm{step}}$ indicates greater similarity between the observed and reference step patterns. Movement efficiency is further measured by the ratio of net displacement to the total trajectory length of both feet. Let $\mathbf{p}_{\mathrm{Lf},t}$ and $\mathbf{p}_{\mathrm{Rf},t}$ denote the left and right foot positions at frame t, respectively. The movement efficiency is defined as

$$\eta = \frac{1}{2}\left(\frac{\left\|\mathbf{p}_{\mathrm{Lf},t_{\mathrm{end}}}-\mathbf{p}_{\mathrm{Lf},t_{\mathrm{sta}}}\right\|}{\sum_{t=t_{\mathrm{sta}}}^{t_{\mathrm{end}}-1}\left\|\mathbf{p}_{\mathrm{Lf},t+1}-\mathbf{p}_{\mathrm{Lf},t}\right\|+\epsilon} + \frac{\left\|\mathbf{p}_{\mathrm{Rf},t_{\mathrm{end}}}-\mathbf{p}_{\mathrm{Rf},t_{\mathrm{sta}}}\right\|}{\sum_{t=t_{\mathrm{sta}}}^{t_{\mathrm{end}}-1}\left\|\mathbf{p}_{\mathrm{Rf},t+1}-\mathbf{p}_{\mathrm{Rf},t}\right\|+\epsilon}\right) \tag{21}$$

where t_{sta} and t_{end} denote the start and end frames of the effective hitting interval, respectively, and ϵ is a small constant that prevents division by zero. The efficiency score is computed separately for the left and right feet and then averaged to account for potential asymmetry in footwork. The final footwork SPI is defined as

$$s_5 = \alpha s_{5,\mathrm{step}} + (1-\alpha)\eta \tag{22}$$

where $\alpha \in [0,1]$ controls the relative contribution of step-count consistency and movement efficiency. Since both $s_{5,\mathrm{step}}$ and η are bounded within $[0,1]$, the resulting index s_5 also lies within this range. Higher s_5 values indicate a footwork pattern closer to the reference, exhibiting higher lower-body movement efficiency.

• SPI-6: Racket control index. This index evaluates the alignment and stability of the racket face near the contact moment. Let $\mathbf{n}_{\mathrm{rac},t}$ and $\mathbf{n}_t^{\mathrm{ref}}$ denote the estimated racket-face and reference normal vectors at frame t, respectively. The angular deviation from the reference orientation is computed as

$$\phi_t = \cos^{-1}\left(\frac{\mathbf{n}_{\mathrm{rac},t}\cdot\mathbf{n}_t^{\mathrm{ref}}}{\left\|\mathbf{n}_{\mathrm{rac},t}\right\|\left\|\mathbf{n}_t^{\mathrm{ref}}\right\|}\right) \tag{23}$$

The alignment score is then defined as

$$s_{6,\mathrm{align}} = \exp\left(-\frac{\bar{\phi}^2}{2\sigma_{\mathrm{align}}^2}\right) \tag{24}$$

where $\bar{\phi}$ denotes the mean angular deviation within the evaluation interval around the contact moment, and $\sigma_{\text{align}} = 0.5$ controls the tolerance to racket-orientation error. To assess racket-face stability, the frame-to-frame angular variation is calculated as

$$\Delta\phi_t = \cos^{-1}\left(\frac{\mathbf{n}_{\text{rac},t} \cdot \mathbf{n}_{\text{rac},t+1}}{\|\mathbf{n}_{\text{rac},t}\| \|\mathbf{n}_{\text{rac},t+1}\|}\right) \tag{25}$$

where the standard deviation of $\Delta\phi_t$ is defined by the stability error within the same evaluation interval: $e_{\text{stab}} = \text{std}(\Delta\phi_t)$. The corresponding stability score is given by

$$s_{6,\text{stab}} = \exp\left(-\frac{e_{\text{stab}}^2}{2\sigma_{\text{stab}}^2}\right) \tag{26}$$

where $\sigma_{\text{stab}} = 0.5$ controls the tolerance to racket face variation. Finally, the racket control index is obtained by combining the alignment and stability terms:

$$s_6 = \beta s_{6,\text{align}} + (1-\beta) s_{6,\text{stab}} \tag{27}$$

where $\beta \in [0,1]$ determines the relative contributions of racket-face alignment and stability. Since both $s_{6,\text{align}}$ and $s_{6,\text{stab}}$ are bounded within $[0,1]$, the resulting score s_6 also lies within this range. Higher s_6 values indicate racket-face orientation that more closely matches the reference pattern and remains more stable near impact.

After obtaining the six normalized quality scores, each score is linearly mapped from the $[0,1]$ interval to a five-point rating scale for visualization and feedback generation: $r_i = 1 + 4s_i$, $i = 1,2,\ldots,6$, where s_i denotes the i-th SPI, and r_i represents the corresponding five-point rating. Ratings closer to 5 indicate stronger performance, whereas ratings near 1 suggest greater need for improvement. The six index-based ratings are visualized on a radar chart using a 1–5 scale (0.5-point tick intervals) to facilitate intuitive interpretation of strengths and weaknesses across different performance dimensions.

The six SPIs were not arbitrarily selected but were developed through discussions with an experienced badminton coach and informed by practical coaching criteria. The proposed SPIs were intentionally selected to represent the biomechanical aspects commonly considered in badminton training. Because no universally accepted reference values exist for the quantitative evaluation of badminton strokes, the reference and tolerance parameters used in the proposed SPI formulation were established through repeated demonstrations by an experienced badminton coach and refined through iterative experimental validation. The effectiveness of the adopted parameters is indirectly validated by the high agreement between the overall stroke-quality assessment and expert evaluations, as presented in Subsection 5.5.

4.2. Hierarchical fuzzy stroke-quality assessment

Although the six SPIs provide interpretable measures of different aspects of stroke execution, badminton stroke quality is inherently a multi-criteria decision-making problem involving complex interactions. Therefore, an HFIS is employed to integrate the six SPIs into an overall stroke-quality assessment, organized into three biomechanically meaningful categories. The first category, termed swing mechanics, evaluates the effectiveness and consistency of the racket swing using the swing velocity index (SPI-1) and wrist trajectory smoothness index (SPI-2). The second category, termed joint coordination, assesses upper-limb posture control through the wrist posture index (SPI-3) and elbow posture index (SPI-4). The third category, termed stroke control, evaluates movement preparation and impact control using the footwork efficiency index (SPI-5) and racket control index (SPI-6). This hierarchical organization follows the biomechanical execution sequence of badminton strokes and simultaneously reduces the complexity of fuzzy rule construction.

4.2.1. Proposed hierarchical fuzzy inference architecture

Figure 10 illustrates the proposed hierarchical FIS for assessing badminton stroke quality. The objective of this design is to integrate multiple SPIs while preserving interpretability and reducing the complexity of fuzzy rule construction. As shown in Figure 10, the first-level inference layer consists of three FIS modules, denoted as F_1, F_2, and F_3. These modules independently evaluate swing quality, joint coordination quality, and stroke control quality. The corresponding outputs are represented as $Q_{\text{swing}} = F_1(s_1, s_2)$, $Q_{\text{joint}} = F_2(s_3, s_4)$, and $Q_{\text{control}} = F_3(s_5, s_6)$. The second-level fuzzy inference layer further integrates the three intermediate assessments to generate the final stroke-quality score, formulated as

$$Q_{\text{stroke}} = F_4\left(Q_{\text{swing}}, Q_{\text{joint}}, Q_{\text{control}}\right) \tag{28}$$

where $Q_{\text{stroke}} \in [0,1]$ denotes the final normalized stroke-quality score generated by the hierarchical FIS.

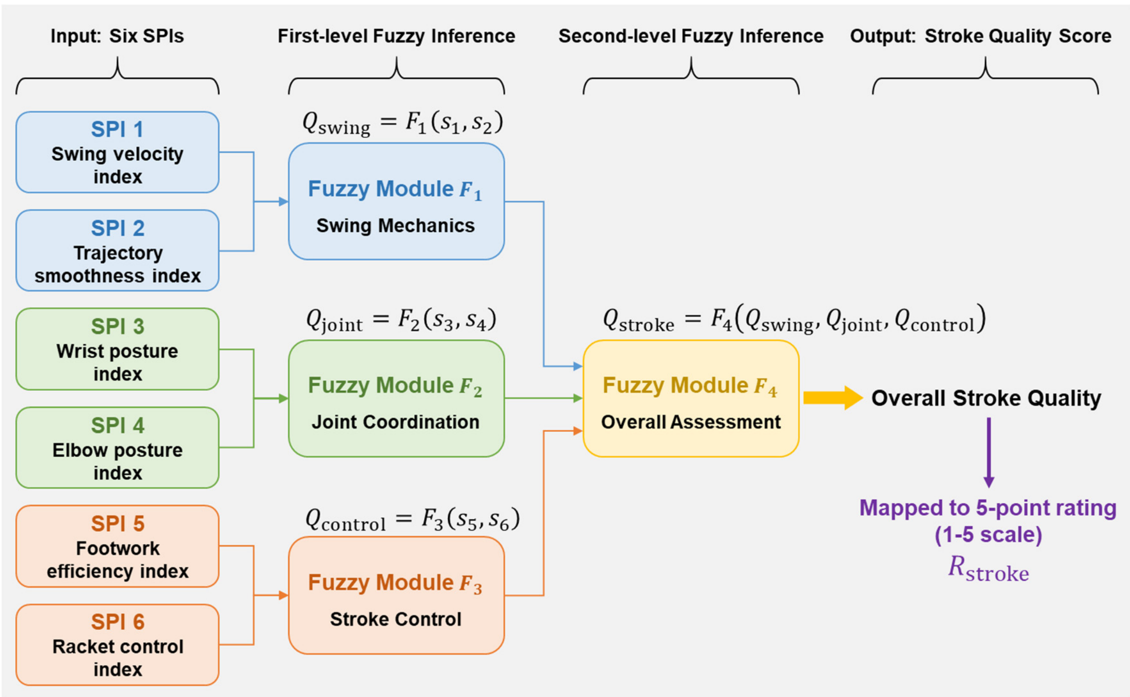


Figure 10. Proposed hierarchical fuzzy inference architecture integrating six SPIs into the final stroke quality score.

4.2.2. Membership functions and rule base

The primary components of implementing an FIS are fuzzification, fuzzy rules, the inference engine, and defuzzification. Let a_1 and a_2 be two premise variables of the FIS, and q denote the corresponding consequent variable. Figures 11 and 12 illustrate the fuzzy sets of the premise and consequent variables, respectively. For both premise variables, trapezoidal and triangular membership functions are adopted to represent the linguistic terms *Low*, *Medium*, and *High*. Similarly, the consequent variable is represented using the linguistic partition: *Very low*, *Low*, *Medium*, *High*, and *Very high*.

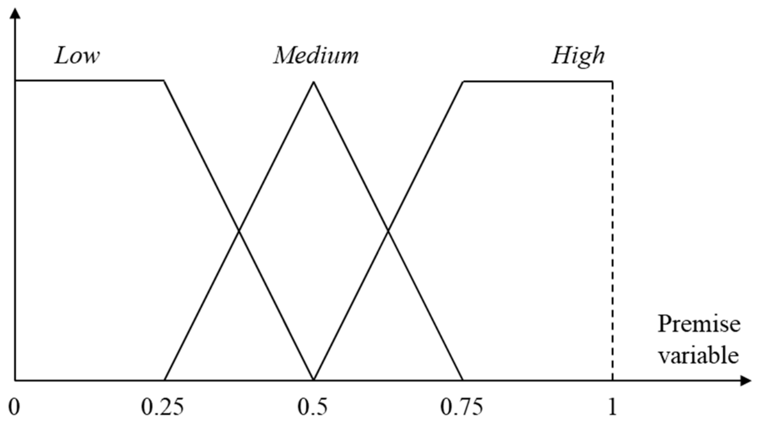


Figure 11. Fuzzy sets of the premise variable.

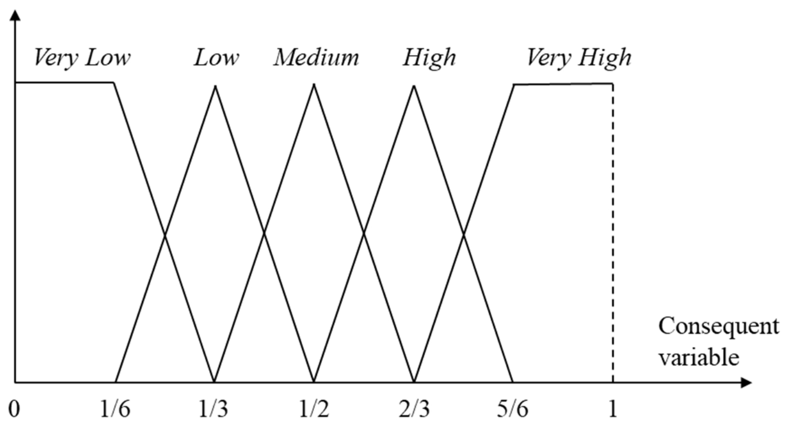


Figure 12. Fuzzy sets of the consequent variable.

Table 1. Fuzzy rule base used in the first-level fuzzy inference modules (F_1, F_2, and F_3).

a_1	a_2	g
Low	*Low*	*Very low*
Low	*Medium*	*Low*
Low	*High*	*Medium*
Medium	*Low*	*Low*
Medium	*Medium*	*Medium*
Medium	*High*	*High*
High	*Low*	*Medium*
High	*Medium*	*High*
High	*High*	*Very high*

Table 2. Fuzzy rule base used in the second-level fuzzy inference module (F_4).

Q_{swing}	Q_{joint}	Q_{control}	Q_{stroke}
Low	*Low*	*Low*	*Very low*
Low	*Low*	*Medium*	*Low*
Low	*Low*	*High*	*Medium*
Low	*Medium*	*Low*	*Low*
Low	*Medium*	*Medium*	*Medium*
Low	*Medium*	*High*	*Medium*
Low	*High*	*Low*	*Low*
Low	*High*	*Medium*	*Medium*
Low	*High*	*High*	*Medium*
Medium	*Low*	*Low*	*Low*
Medium	*Low*	*Medium*	*Medium*
Medium	*Low*	*High*	*Medium*
Medium	*Medium*	*Low*	*Medium*
Medium	*Medium*	*Medium*	*Medium*
Medium	*Medium*	*High*	*Medium*
Medium	*High*	*Low*	*Medium*
Medium	*High*	*Medium*	*Medium*
Medium	*High*	*High*	*High*
High	*Low*	*Low*	*Medium*
High	*Low*	*Medium*	*Medium*
High	*Low*	*High*	*High*
High	*Medium*	*Low*	*Medium*
High	*Medium*	*Medium*	*Medium*
High	*Medium*	*High*	*High*
High	*High*	*Low*	*Medium*
High	*High*	*Medium*	*High*
High	*High*	*High*	*Very high*

For the swing-mechanics module F_1, the premise variables are defined as $a_1 = s_1$ and $a_2 = s_2$, where s_1 and s_2 denote the swing velocity index and wrist trajectory smoothness index, respectively, and the consequent variable is the swing-quality score Q_{swing}. Similarly, for the joint-coordination module F_2, $a_1 = s_3$ and $a_2 = s_4$, where s_3 and s_4 denote the wrist posture index and elbow posture index, respectively, and the consequent variable is the joint-coordination score Q_{joint}. For the stroke-control module F_3, $a_1 = s_5$ and $a_2 = s_6$, where s_5 and s_6 denote the footwork efficiency index and racket control index, respectively, and the consequent variable is the stroke-control score Q_{control}. At the second inference level, the three intermediate assessments Q_{swing}, Q_{joint}, and Q_{control} serve as the premise variables of the fuzzy module F_4, while the consequent variable is the final stroke-quality score Q_{stroke}. The fuzzy rules of the first-level inference layer are subsequently constructed using the defined linguistic variables. For example,

IF a_1 is *High* **AND** a_2 is *High*, **THEN** g is *Very high*.

Following this principle, the fuzzy rule base can be established objectively. Table 1 shows the rule table used in F_1, F_2, and F_3. The fuzzy rule base adopted in the second-level inference module F_4 is summarized in Table 2. Some combinations of linguistic conditions may be less likely to occur in real badminton strokes because the biomechanical components of stroke execution are not entirely independent. Nevertheless, these rules are retained to guarantee completeness of the fuzzy rule base and to avoid undefined inference regions.

4.2.3. Mamdani inference and defuzzification

As an input tuple $(\tilde{a}_1, \tilde{a}_2)$ goes into the FIS, it activates the corresponding fuzzy rules according to their membership degrees. The proposed hierarchical fuzzy inference system employs the Mamdani inference mechanism, in which the logical "AND" operation is implemented using the minimum operator. For a generic fuzzy rule,

$$\textbf{IF } a_1 \text{ is } A_1^m \textbf{ AND } a_2 \text{ is } A_2^m, \textbf{THEN } g \text{ is } G^m,$$

the firing strength of the m-th rule is computed as

$$\mu_m = \min\left(A_1^m(\tilde{a}_1), A_2^m(\tilde{a}_2)\right) \tag{29}$$

where $A_1^m(\cdot)$ and $A_2^m(\cdot)$ denote the membership functions associated with the premise variables. The aggregated consequent fuzzy set is then obtained as

$$G'(g) = \max_m \min\left(\mu_m, G^m(g)\right) \tag{30}$$

where $G^m(g)$ denotes the consequent fuzzy set of the m-th rule. Following rule aggregation, the center-of-gravity defuzzification method [56] is employed to obtain a crisp output value. The crisp output is calculated as

$$g^* = \left(\int_{\mathbb{G}} g \cdot G'(g)dg\right) / \left(\int_{\mathbb{G}} G'(g)dg\right) \tag{31}$$

where $\mathbb{G}$ denotes the universe of discourse of the output variable. For the second-level inference module F_4, the same inference mechanism is adopted. Therefore, the corresponding firing-strength computation, fuzzy aggregation, and centroid defuzzification can be formulated analogously to Eqs (29)–(31).

Through the first- and second-level inference processes, the resulting value $Q_{\text{stroke}} \in [0,1]$ is obtained, where larger values indicate better overall stroke quality. For visualization and LLM-assisted feedback generation, the final assessment score is mapped to a five-point rating scale as

$$R_{\text{stroke}} = 1 + 4Q_{\text{stroke}} \tag{32}$$

This overall stroke-quality rating (R_{stroke}) also serves as a practical reference for identifying effective strokes. Based on our experimental observations, most effective strokes achieved overall scores above 3.5, although this threshold was not optimized and may be further refined in future work. Therefore, the overall score provides an intuitive quantitative reference for stroke evaluation and supports subsequent LLM-assisted coaching analytics.

4.3. LLM-assisted coaching analytics

4.3.1. Structured inputs and outputs of the LLM

Although the proposed scoring framework provides quantitative stroke-quality assessment through the SPIs and hierarchical fuzzy inference system, numerical scores alone may be difficult for general users to interpret and translate into actionable training strategies. Therefore, the proposed framework employs Google Gemini 2.5 Pro as the LLM engine to transform structured assessment results into personalized natural-language coaching feedback. In the proposed framework, the LLM receives structured inputs from the preceding modules, including the recognized stroke type, the six SPIs, the intermediate fuzzy assessment scores (Q_{swing}, Q_{joint}, and Q_{control}), the final stroke-quality score (Q_{stroke}), and the corresponding radar-chart ratings, ensuring feedback is grounded in quantitative motion analysis and fuzzy assessment results.

Based on the structured inputs, the LLM generates a comprehensive coaching report consisting of 1) an overall performance summary, 2) an analysis of technical strengths and deficiencies, and 3) personalized training recommendations. For example, a high swing-mechanics assessment score combined with a low stroke-control score may indicate adequate racket acceleration but insufficient footwork preparation or racket-face stability. Similarly, a low joint-coordination score may suggest suboptimal wrist or elbow posture during stroke execution.

The proposed framework accesses Gemini 2.5 Pro through Google Gemini API. Unless otherwise specified, the API is used with its default generation settings, while the domain-specific behavior of the LLM is controlled through the structured prompt construction described in the following subsection.

4.3.2. Knowledge-guided structured prompting strategy

To enhance consistency and reliability, a knowledge-guided structured prompting strategy is employed to define the LLM's role, response scope, and output format. The prompt constrains the LLM to generate feedback solely from the recognized stroke type and the quantitative assessment results produced by the proposed framework, without inferring player skill level, physical condition, or tactical intentions. To further improve the technical consistency of the generated feedback, a badminton coaching knowledge base was manually constructed. The knowledge was drawn from badminton training books and refined through discussion with an experienced badminton coach. During the prompting process, the knowledge was organized into structured mappings linking the six SPIs to stroke-quality interpretations and training recommendations. This manually curated knowledge base was incorporated into prompt construction before LLM inference.

Furthermore, a human-in-the-loop verification mechanism is incorporated to validate the quality of the generated coaching feedback. The evaluation panel consists of one badminton coach, one physical education instructor, and two badminton athletes. They assess whether the generated responses are technically accurate, consistent with the SPI representation and fuzzy assessment results, and appropriate for the recognized stroke type. This design bridges explainable computational intelligence and practical coaching communication, producing interpretable and actionable training feedback.

5. Experiments and discussion

5.1. Experimental setup and dataset

To evaluate the proposed framework, experiments were conducted to assess its main functional modules, including stroke event segmentation, pose estimation, motion feature extraction, stroke type recognition, quality scoring, and LLM-assisted feedback generation. The experiments verified that the proposed pipeline transforms badminton stroke videos into interpretable SPIs, stroke-level quality scores, and natural-language coaching feedback.

Table 3. Stroke categories used for stroke type recognition.

Class ID	Stroke category	Description
1	Forehand clear	A forehand overhead stroke that sends the shuttlecock deep to the opponent's backcourt.
2	Forehand cut	A forehand stroke with a slicing motion to change the shuttlecock speed or trajectory.
3	Forehand drive	A fast and flat forehand stroke directed toward the opponent's midcourt or rear court.
4	Forehand drop	A controlled forehand stroke that sends the shuttlecock softly toward the front court.
5	Forehand lift	A forehand underhand stroke that lifts the shuttlecock toward the opponent's backcourt.
6	Backhand clear	A backhand stroke that sends the shuttlecock deep to the opponent's backcourt.
7	Backhand cut	A backhand slicing stroke used to alter the shuttlecock's speed or direction.
8	Backhand drive	A fast and flat backhand stroke directed across the court.
9	Backhand drop	A controlled backhand stroke that sends the shuttlecock softly toward the front court.
10	Backhand lift	A backhand underhand stroke that lifts the shuttlecock toward the opponent's backcourt.
11	Smash	A powerful overhead attacking stroke directed downward toward the opponent's court.

The experimental data were collected in two settings. For stroke type recognition, a stereo-vision system was used to acquire training and test videos for recognition dataset construction and evaluation of the proposed mST-GCN++ model. The dataset was collected from six participants performing badminton strokes under coach-fed conditions. The original recordings included multiple repetitions of 12 stroke categories. After video segmentation and preprocessing, the final recognition dataset comprised 3126 segmented stroke clips, including 2490 training and validation samples and 636 test samples. For the recognition task in this study, the left-smash and right-smash classes were merged into a single smash category, resulting in 11 stroke categories, as summarized in Table 3. A summary of the dataset statistics used for stroke recognition is provided in Table 4. The distribution of stroke clips across the 11 stroke categories is summarized in Table 5. Each participant contributed 521 stroke clips, including 51 smash clips and 47 clips for each of the remaining 10 stroke categories, resulting in a total of 3126 stroke clips with a nearly balanced class distribution.

Second, stroke videos recorded using a smartphone application in real training scenarios were uploaded to a cloud-based analysis system. The resulting radar-chart scores and AI-generated coaching suggestions were then returned to the mobile application for user review, enabling evaluation of the proposed framework under deployment-oriented conditions using accessible recording devices.

To summarize, the stereo-vision setup was primarily used for reliable dataset construction and stroke recognition model development, whereas the smartphone application was used to validate the practical assessment and feedback workflow. This experimental design integrates controlled model development with real-world usability evaluation.

Table 4. Dataset statistics for stroke recognition.

Item	Value
Participants	6
Stroke categories (as listed in Table 3)	11
Clips per participant	521
Total clips	3126
Input representation	Skeletal motion sequences
Data acquisition	Stereo-vision system

Table 5. Distribution of the stroke recognition dataset across the 11 stroke categories.

Stroke type	Clips per participant	Clips
Forehand clear	47	282
Forehand cut	47	282
Forehand drive	47	282
Forehand drop	47	282
Forehand lift	47	282
Backhand clear	47	282
Backhand cut	47	282
Backhand drive	47	282
Backhand drop	47	282
Backhand lift	47	282
Smash	51	306
Total	521	3126

5.2. Stroke event segmentation results

Stroke event segmentation was evaluated to assess accurate detection of racket–shuttlecock contact moments and stroke clip extraction from continuous videos. Reliable contact localization is critical, as subsequent modules rely on correctly segmented clips. In this study, a YOLOv11s-based object detector [57] was trained to identify the shuttlecock and racket in each frame, chosen for its small model size and low computational complexity, making it suitable for real-time applications. The dataset comprised 10,615 training and 1875 test images (collected using our stereo vision system and a smartphone). Training used a batch size of 16, 640×640 pixel resolution, and a maximum of 500 epochs, with early stopping (patience 20).

Following Subsection 3.1, contact moments were determined from the normalized racket–shuttlecock centers' spatial distance falling below 0.1, treated as candidate contact frames. The frame with the minimum distance was selected as the final hitting moment. Compared to manual annotations (the ground-truth contact events in Table 6), the YOLOv11s-based method detected most hitting moments, with few missed or false positives. The resulting precision, recall, and F1-score confirm reliable localization of contact moments in badminton stroke videos. A detection was considered correct when the estimated contact frame differed from the manually annotated frame by no more than ± 1 frame.

Table 6. Quantitative evaluation of the contact-frame detection on the test dataset.

Item	Value	Description
Ground-truth contact events	1373	Total manually annotated racket–shuttlecock contact events in the test dataset.
Correctly detected contact events (TP)	1368	Contact events correctly identified by the method.
Missed contact events (FN)	5	Ground-truth contact events that were not detected.
False contact detections (FP)	2	Non-contact frames incorrectly identified as contact events.
Precision	99.85%	Ratio of correctly detected contact events to all detected contact events, $\mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$
Recall	99.64%	Ratio of correctly detected contact events to all ground-truth contact events, $\mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$
F1-score	99.74%	Harmonic mean of precision and recall.

5.3. Stroke type recognition results

5.3.1. Overall recognition performance

Stroke type recognition was evaluated to assess the correct classification of segmented badminton strokes. The recognized stroke type provides essential contextual information for motion interpretation, quality scoring, and LLM-assisted feedback generation. As summarized in Tables 4 and 5, the recognition dataset consisted of 3126 samples collected from six participants. For overall recognition evaluation, the dataset was randomly divided into 2490 training and validation samples (approximately 80%) and 636 testing samples (approximately 20%), while maintaining the class distribution across all stroke categories. The segmented stroke clips were transformed into skeletal motion sequences and then input into the recognition model across 11 strokes. As summarized in Table 7, the proposed mST-GCN++ achieved the highest accuracy of 94.34%, outperforming baseline ST-GCN++. This result indicates that integrating multi-scale temporal convolution with Transformer-based temporal modeling enhances badminton stroke recognition performance.

Table 7. Numerical comparison of stroke type recognition accuracy.

Model	Role	Accuracy
CTR-GCN [32]	Baseline	91.35%
MS-AAGCN [28]	Baseline	92.30%
ST-GCN++ [53]	Baseline backbone	92.92%
ST-GCN++ with Modified MS-TCN	Modified variant	93.87%
ST-GCN++ with PatchTST	Modified variant	93.55%
mST-GCN++ (ours)	**Final recognition model**	**94.34%**

5.3.2. Subject-independent validation using LOSO

Because the above evaluation used a random train-test split, this protocol may allow data from the same participant to appear in both the training and test datasets, potentially overestimating the model's generalization capability. To provide a more rigorous evaluation, a leave-one-subject-out (LOSO) cross-validation protocol was also conducted. In each validation fold, all 521 stroke clips belonging to one participant were reserved exclusively for testing, while the remaining 2605 stroke clips from the other five participants were used for model training. Consequently, each participant served as the test subject once, resulting in six independent validation folds with no subject overlap between the training and test datasets. The recognition performance for each validation fold is summarized in Table 8.

Table 8. Subject-independent stroke-recognition performance of the proposed mST-GCN++ using LOSO cross-validation.

Fold	Subject	Accuracy	Precision	Recall	F1-score
1	S1	0.9060	0.9149	0.9061	0.9105
2	S2	0.8983	0.9044	0.8982	0.9013
3	S3	0.9136	0.9151	0.9136	0.9143
4	S4	0.9175	0.9240	0.9173	0.9206
5	S5	0.9098	0.9133	0.9098	0.9115
6	S6	0.9251	0.9267	0.9250	0.9258
Mean ± SD		0.9117 ± 0.00849	0.9164 ± 0.00732	0.9117 ± 0.00846	0.9140 ± 0.00779

Note: Precision, recall, and F1-score denote macro-averaged metrics over the 11 stroke categories. All performance metrics are reported as decimal values rather than percentages.

The proposed mST-GCN++ achieved an average recognition accuracy of 91.17%, demonstrating stable performance across different participants. As expected, the LOSO accuracy is slightly lower than the accuracy obtained with the random train-test split, since the model is evaluated on previously unseen players. Nevertheless, the consistently high recognition performance indicates that the proposed framework effectively captures stroke-specific motion patterns rather than subject-specific characteristics, demonstrating good subject-independent generalization. Although LOSO validation demonstrates improved subject generalization, the current dataset was collected from a limited number of participants under relatively controlled recording conditions. Future work will further evaluate the proposed framework using larger-scale datasets involving players with diverse skill levels, playing styles, and recording environments.

5.4. Stroke-quality assessment and feedback results

The stroke-quality assessment and feedback module was evaluated using three representative stroke cases, as described below, to demonstrate the feasibility and interpretability of the proposed pipeline, rather than through large-scale statistical validation. For both the five-point indicator ratings and the overall stroke-quality score, a score of 3.5 (on a five-point scale) was used as the reference threshold for determining effective strokes. Scores below this threshold indicate room for technical improvement.

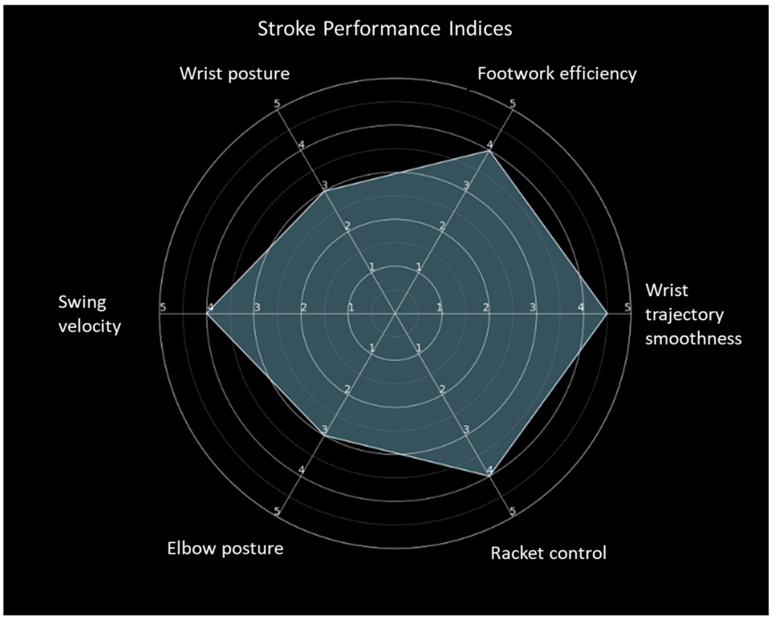


Figure 13. Radar-chart visualization of the stroke-quality assessment results for Case 1.

• Effective smash stroke: Figure 13 presents the radar chart for an effective smash stroke. All six SPIs received ratings above the level of 3.0, with an average indicator rating of 3.75 and an overall stroke-quality rating of 3.63, indicating satisfactory performance across all evaluated aspects. Footwork trajectory, swing velocity, wrist trajectory smoothness, and racket control indices reached approximately 4 points, reflecting adequate swing acceleration, stable racket-arm motion, and effective lower-body positioning. The wrist angle remained mostly within the expected range, with only a minor mid-swing deviation that quickly returned to the reference range. Figure 14 presents the LLM-assisted coaching feedback: smooth lower-body positioning, stable footwork, good wrist and elbow angle control, and stable racket face control. Despite a high swing velocity rating, the feedback suggested that excessively fast execution should be balanced with timing and racket-face control. Accordingly, it recommends focusing on hitting timing, practicing controlled swings, and progressively increasing swing velocity to optimize power, accuracy, and control.

• Effective forehand drop stroke: Figure 15 illustrates the effective forehand drop stroke. Most indicators met the acceptable level, though the elbow joint angle was slightly lower than the remaining indicators. This result suggests successful stroke execution, but with room for arm extension and impact posture. Wrist trajectory remained on a generally stable swing path, but the hand position exhibited a slight forward movement compared with the reference motion. Figure 16 presents the LLM-assisted coaching feedback. Due to space limitations, only the overall analysis and targeted training suggestions are reported. The feedback indicates that the player demonstrated stable lower-body positioning and good racket-face control, but improved distance judgment and hitting timing are needed. The feedback further recommends footwork practice and hitting-feeling practice to enhance movement efficiency, stroke timing, and shot placement when executing forehand drop shots.

1. Wrist Trajectory Smoothness:
After hitting the shuttlecock, a slight lateral deviation of the wrist was observed. It is recommended to maintain wrist stability at the moment of impact and avoid excessive rotation or deviation. The player should practice maintaining the stability of the wrist centerline and try to return the wrist quickly to its initial position after impact. Slow-motion practice can be used to observe the wrist rotation trajectory.

2. Wrist Posture:
The wrist angle control is acceptable and should be maintained. The player is encouraged to continue maintaining a proper wrist angle, as this helps improve hitting accuracy and shot control.

3. Swing Velocity:
The swing speed is slightly faster than the standard level, and the player should pay attention to the hitting timing. It is recommended to practice controlling the swing speed and finding the optimal hitting moment. The best hitting performance can be achieved only with an appropriate swing speed. The player may practice slow swings first and then gradually increase the speed to find the best balance. During practice, the focus should be placed on hitting accuracy and control rather than simply pursuing speed.

4. Elbow Posture:
The elbow angle control is acceptable and should be maintained. Maintaining a proper elbow angle helps improve swing efficiency and force transmission.

5. Footwork Efficiency:
No obvious deviation was observed in the trajectories of both feet. The player is encouraged to continue maintaining a stable footwork trajectory, which is essential for maintaining balance and improving reaction speed.

6. Racket Control:
The racket face control is acceptable and should be maintained. Maintaining accurate racket face control is highly beneficial for improving hitting accuracy.

Overall Analysis:
Based on the analysis of the six major aspects, as well as the good performance in body center-of-mass transfer, a step count of three steps, and a movement speed of 1.1 m/s, the overall performance of this smash stroke was quite good.

Strengths:
The body center-of-mass transfer was smooth, the footwork was stable, the wrist and elbow angle control were good, and the racket face control was also stable.

Weakness:
The swing speed was slightly faster than the standard level, which may affect hitting accuracy.

Targeted Training Suggestions:
It is recommended to focus on mastering the hitting timing, practice slow swings, and gradually increase the swing speed to find the best balance between swing speed and hitting accuracy. The player may also practice training exercises that improve swing control, such as adding resistance during the swing to enhance stability and accuracy. The player should continue maintaining good body center-of-mass transfer and stable footwork.

Figure 14. Example of LLM-assisted coaching feedback for Case 1 (full response).

• Unsuccessful forehand clear stroke: Figure 17 shows the third case, an unsuccessful forehand clear stroke. Compared with the previous successful cases, this stroke exhibits reduced stability across multiple SPIs. The radar-chart result indicates that swing control and wrist trajectory are less consistent throughout execution. In particular, the swing velocity exceeded the reference level, suggesting excessive force application. The wrist trajectory also shows greater deviation, with the hand moving outward during the preparation phase and overextending forward near impact. Figure 18 presents the corresponding LLM-assisted coaching feedback. Due to space limitations, only the overall analysis and targeted training suggestions are reported. The unsuccessful stroke involved excessive backswing, unstable wrist trajectory, and improper force control. Recommendations include reducing unnecessary arm extension, improving wrist stability, and practicing smoother swing timing to enhance stroke accuracy and consistency.

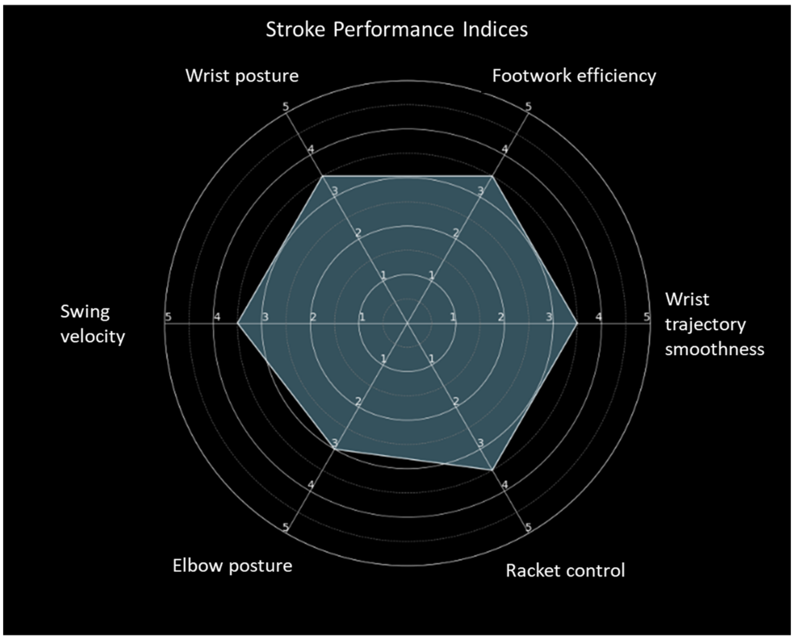


Figure 15. Radar-chart visualization of the stroke-quality assessment results for Case 2.

Overall Analysis:
Based on the analysis of the six major aspects, the player showed good center-of-mass transfer, a step count of two steps, and a movement speed of 0.9 m/s. The results indicate that, during this forehand drop stroke, the player should pay attention to distance judgment to avoid missing the optimal hitting timing. The center-of-mass transfer and racket face control were performed well and should be maintained.
Targeted Training Suggestions:
Footwork Practice:
Focus on correct positioning and footwork movement. During practice, the player can observe their own movement using a mirror or ask a coach for guidance. The player can also practice faster footwork movements and combine them with different hitting positions to improve movement efficiency on the court.
Hitting Feeling Practice:
Practice different hitting sensations and hitting strategies. Hitting technique, shot direction, and coordination with court positioning are all key factors for improving stroke performance.

Figure 16. Example of LLM-assisted coaching feedback for Case 2 (excerpted response).

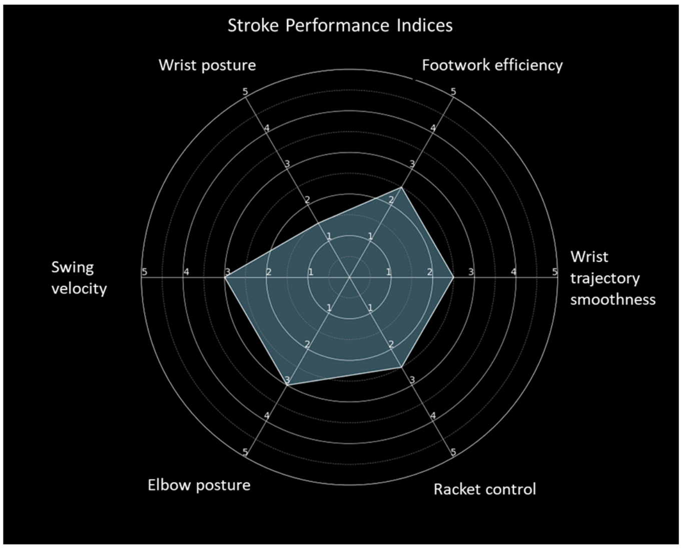


Figure 17. Radar-chart visualization of the stroke-quality assessment results for Case 3.

Overall Analysis:
Based on the analysis of the six major aspects, the player showed unstable wrist trajectory, excessive swing speed, and insufficient control of the hitting posture during this forehand clear stroke. The results indicate that the unsuccessful return was mainly related to excessive backswing, inappropriate force control, and poor timing near the impact moment. The player should focus on maintaining a more stable swing path and avoiding unnecessary forward extension of the hitting arm.
Targeted Training Suggestions:
Swing Control Practice:
Focus on controlling the swing range and swing speed. During practice, the player should avoid using excessive force and instead emphasize smooth acceleration and stable racket control. The player can begin with slow and controlled forehand clear swings, then gradually increase the swing speed while maintaining accuracy and consistency.
Wrist Trajectory Practice:
Practice maintaining a stable wrist path during the preparation, hitting, and follow-through phases. The player should avoid moving the wrist too far outward during the backswing or extending the hand excessively forward near impact. Slow-motion drills or video feedback can be used to observe the wrist trajectory and correct unnecessary deviations.
Hitting Timing Practice:
Focus on identifying the optimal hitting position and impact timing. The player should practice contacting the shuttlecock at a stable and comfortable position rather than compensating with excessive arm extension. Repeated forehand clear drills with different shuttlecock placements can help improve timing, distance judgment, and stroke consistency.

Figure 18. Example of LLM-assisted coaching feedback for Case 3 (excerpted response).

5.5. Validation of stroke-quality assessment

5.5.1. Human expert evaluation protocol

To quantitatively evaluate the proposed stroke-quality assessment framework under practical deployment conditions, a total of 60 badminton stroke clips were newly collected using a single smartphone camera. These clips were independent of the datasets used for training the shuttlecock detector and stroke-recognition model, thereby providing an unbiased evaluation of the complete assessment framework in real-world scenarios. This dataset covers 11 badminton stroke types and includes both effective and ineffective strokes to ensure diverse stroke qualities and performance levels.

Table 9. Expert rating criteria for overall stroke-quality assessment.

Score	Rating	Description
1.0	Very poor	The stroke is incorrectly executed, lacks coordination and control, and is ineffective in practical play.
2.0	Poor	Multiple technical deficiencies are present, leading to unstable execution and limited effectiveness.
3.0	Acceptable	The stroke can be completed successfully but exhibits noticeable deficiencies in one or more technical aspects, resulting in reduced performance.
4.0	Good	Minor technical imperfections are evident, but the stroke is stable, well-coordinated, and effective in practical play.
5.0	Excellent	The stroke demonstrates excellent coordination, stable posture, accurate racket control, and efficient footwork, and is highly effective in both offensive and defensive situations.

Each clip was independently evaluated by two experienced badminton experts, including one physical education instructor (Expert A) and one badminton coach (Expert B). The experts assessed the overall stroke quality on a five-level semantic rating scale, where 1, 2, 3, 4, and 5 correspond to *very poor*, *poor*, *acceptable*, *good*, and *excellent*, respectively. Intermediate scores (i.e., 1.5, 2.5, 3.5, and 4.5) were permitted when the experts considered the stroke quality to fall between two adjacent performance levels. The semantic definitions associated with each rating level are summarized in Table 9. To assess consistency between the two experts, inter-rater reliability was first evaluated using the intraclass correlation coefficient (ICC). The ICC(2,1) and ICC(3,1) values were 0.8584 and 0.8588, respectively, indicating good agreement between the two experts. After confirming satisfactory agreement, the average of the two expert ratings was used as the expert-rated reference score for each stroke in the subsequent quantitative evaluation in Subsection 5.5.2.

5.5.2. Statistical analysis of stroke-quality assessment

To quantitatively evaluate the proposed stroke-quality assessment framework, the predicted stroke-quality scores were compared with the expert-rated reference scores (established in Subsection 5.5.1) using multiple statistical measures, including the Pearson correlation coefficient, Spearman rank correlation coefficient, intraclass correlation coefficients (ICC(2,1) and ICC(3,1)), mean absolute error (MAE), and root mean square error (RMSE). Here, the reported ICC(2,1) and ICC(3,1) values quantify the agreement between the model-generated stroke-quality scores and the expert-rated reference scores. Three assessment strategies were investigated for comparison: mean, in which the six SPIs were directly averaged without fuzzy inference; HFIS, which employed manually designed membership functions (as plotted in Figures 11 and 12) and fuzzy rules of the hierarchical FIS; and the proposed refined-HFIS, in which the membership-function parameters were manually refined through a local parameter search around the initially specified membership-function parameters while preserving the original hierarchical fuzzy inference structure and fuzzy rule base. Specifically, the refinement process was performed by allowing the parameters of all membership functions to vary within $\pm 20\%$ of the manually designed membership functions. The best-performing parameter combination within this bounded search region was selected to construct the refined-HFIS. The quantitative comparison among the three methods is summarized in Table 10.

Notably, this refinement was conducted using the same 60 expert-rated stroke clips that were later used for performance evaluation. Therefore, the refined-HFIS results should be regarded as an exploratory in-sample refinement intended to assess the sensitivity of the manually designed membership functions, not as an independent out-of-sample validation. Consequently, the original HFIS serves as the primary evidence supporting the effectiveness of the proposed stroke-quality assessment framework, whereas the refined-HFIS is included to illustrate the potential improvement achievable through local parameter refinement. Moreover, the comparison between HFIS and refined-HFIS also implicitly provides a local robustness (or sensitivity) evaluation of the manually designed fuzzy system because the refinement was intentionally restricted to a bounded local neighborhood of the initially specified parameters. If only limited performance improvements are observed after refinement, the original HFIS can be considered relatively insensitive to moderate parameter perturbations, indicating that the manually designed fuzzy model provides a stable stroke-quality assessment.

Table 10. Quantitative comparison of three stroke-quality assessment strategies against expert-rated scores.

Strategy	Pearson↑	Spearman↑	ICC(2,1)↑	ICC(3,1)↑	MAE↓	RMSE↓
Mean	0.676	0.7157	0.6614	0.6588	0.5727	0.6981
HFIS	0.7906	0.8169	0.7878	0.7850	0.4308	0.5185
Proposed refined-HFIS	0.8186	0.8372	0.8174	0.8153	0.3901	0.4736

As summarized in Table 10, directly averaging the six SPIs resulted in the lowest agreement with expert-rated stroke quality, achieving the lowest Pearson correlation, Spearman correlation, ICC values, and the largest prediction errors (MAE and RMSE). Incorporating the hierarchical fuzzy inference system substantially improved agreement with expert evaluations. Finally, the proposed refined-HFIS framework achieved the best performance across all evaluation metrics. These results demonstrate that tuning the membership-function parameters enables the proposed framework to better approximate expert evaluation behavior while maintaining the interpretability of the hierarchical fuzzy inference system. Overall, the experimental results indicate that the proposed HFIS improves the consistency between automatically generated stroke-quality scores and expert coaching assessments. Compared with end-to-end score regression approaches, the proposed framework explicitly separates motion representation (SPIs), reasoning (HFIS), and coaching analytics (LLM). This hierarchical design provides transparent reasoning and enables every coaching recommendation to be traced back to quantitative biomechanical indicators, making the proposed framework more suitable for explainable AI-assisted sports coaching. Although refined-HFIS achieved the highest agreement with expert ratings across all evaluation metrics, the performance improvement over the original HFIS remained modest. This result suggests that the proposed manually designed HFIS is reasonably robust and not overly sensitive to moderate variations in the membership-function parameters, thereby supporting the stability of the proposed fuzzy assessment framework.

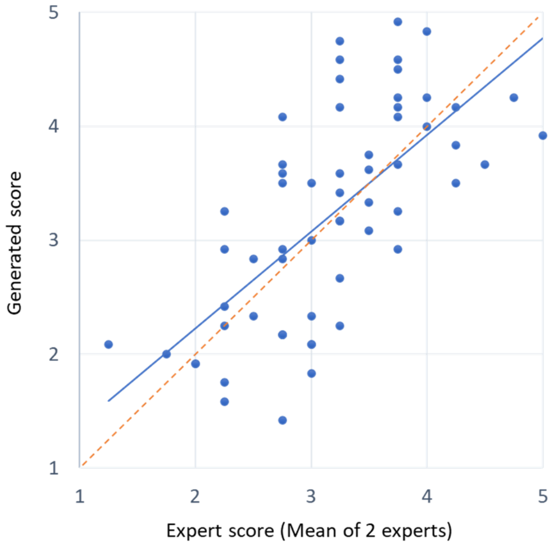


Figure 19. Scatterplot of expert-rated stroke-quality scores versus predicted scores obtained using direct SPI averaging (mean).

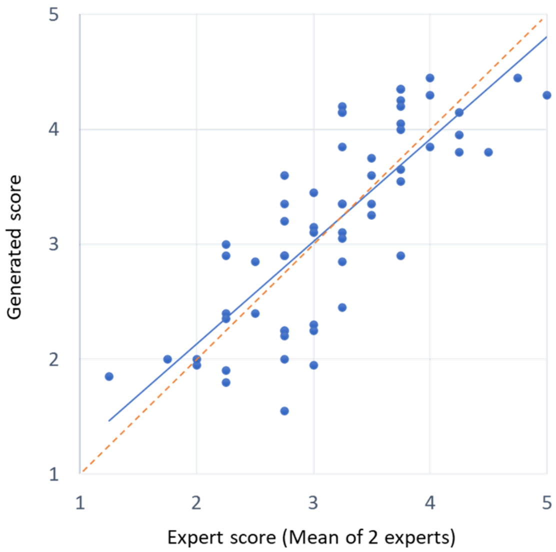


Figure 20. Scatterplot of expert-rated stroke-quality scores versus predicted scores obtained using HFIS.

Moreover, Figures 19–21 provide a visual comparison of the three assessment strategies. In each figure, the blue solid line represents the linear regression fitted to the experimental data, while the orange dashed line denotes the ideal agreement line, where the predicted score exactly matches the expert-rated score. As shown in Figure 19, the direct averaging strategy exhibits relatively large deviations from the ideal agreement line, indicating weaker consistency with expert evaluations. Incorporating the HFIS substantially reduces the score dispersion (Figure 20), suggesting that fuzzy reasoning more effectively models the relationships among multiple stroke performance indices. Finally, the proposed refined-HFIS (Figure 21) exhibits the highest concentration of data points around the ideal agreement line, demonstrating the closest agreement with expert-rated stroke quality.

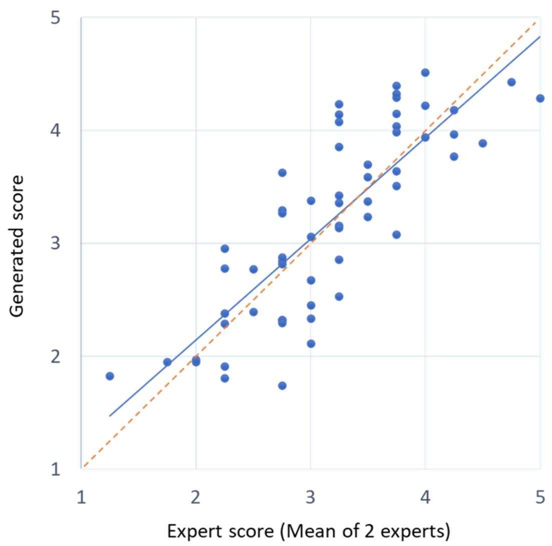


Figure 21. Scatterplot of expert-rated stroke-quality scores versus predicted scores obtained using refined-HFIS.

In addition to the stroke-quality assessment via the refined-HFIS, the practical capability of the proposed framework to distinguish effective and ineffective strokes was also investigated. For the predicted labels, an empirical threshold of 3.5, selected based on experimental observations, was adopted. Refined-HFIS ratings of 3.5 or higher were classified as effective, whereas ratings below 3.5 were classified as ineffective. Independently, the ground-truth labels were obtained. For the ground-truth label, each stroke was manually annotated as either effective or ineffective based on the actual match outcome, as described below. An effective stroke was defined as one that directly resulted in a point, induced an opponent's error, or prevented the opponent from scoring on the immediately subsequent return. In contrast, an ineffective stroke was defined as one that resulted in an immediate unforced error by the player or allowed the opponent to score directly on the next return. These manually assigned labels reflect not only the technical quality of a stroke but also the subsequent rally outcome, which may be influenced by the opponent's response and tactical context. The corresponding classification performance is summarized in Table 11. From this table, the proposed refined-HFIS achieved an overall classification accuracy of 0.9167 in distinguishing effective and ineffective badminton strokes, with a precision of 0.8750, recall of 0.9130, and F1-score of 0.8936. These results suggest that the proposed stroke-quality assessment is generally consistent with practical rally outcomes under the adopted evaluation protocol. However, because the effective/ineffective labels are also influenced by the opponent's response and tactical context, this analysis should be regarded as supplementary exploratory evidence rather than an independent validation of technical stroke quality.

Table 11. Binary classification results for distinguishing effective and ineffective strokes.

	Predicted class	
	Effective	Ineffective
Actual class		
Effective	21	2
Ineffective	3	34

5.6. Human-in-the-loop feedback verification

Table 12. Expert evaluation questionnaire and assessment results for the proposed LLM-assisted badminton stroke feedback system.

Evaluation criterion	Expert A	Expert B	User C	User D	Mean score
Technical correctness					
1. The generated feedback accurately reflects the player's stroke performance.	8.43	8.50	8.77	8.80	8.625
2. The feedback is technically correct from a badminton coaching perspective.	8.60	8.47	8.73	8.77	8.643
3. The system correctly identifies important movement characteristics during stroke execution.	8.73	8.57	8.33	8.40	8.508
Motion consistency and interpretability					
4. The generated feedback is consistent with the observed body movements in the video.	8.40	8.40	8.93	8.90	8.658
5. The SPIs and assessment results are easy to understand and interpret.	8.63	8.50	8.53	8.40	8.515
6. The system provides reasonable explanations for the assigned scores and evaluations.	8.80	8.67	8.67	8.73	8.718
Practical coaching usefulness					
7. The generated suggestions are useful for practical badminton training.	8.57	8.60	8.60	8.77	8.635
8. The feedback provides meaningful guidance for improving stroke technique.	8.30	8.37	8.57	8.43	8.418
9. The system could assist coaches or physical education instructors during training sessions.	8.60	8.50	8.73	8.67	8.625
Overall system evaluation					
10. The proposed system demonstrates good practical usability and coaching potential.	8.53	8.60	8.67	8.60	8.600
Overall mean score	**8.559**	**8.518**	**8.653**	**8.647**	**8.594**

Note: Scores are reported on a 10-point scale and averaged across 30 evaluated stroke videos. Expert A represents a physical education instructor, whereas Expert B represents an experienced badminton coach. Users C and D are badminton athletes who participated in the evaluation to provide athlete-oriented perspectives.

To evaluate the practical usability of the LLM-assisted feedback module, two domain experts and two badminton athletes were invited to assess 30 stroke clips (sampled from the 60 clips in Subsection 5.5) along with the corresponding AI-generated training suggestions. For each case, the experts reviewed the video, system-generated assessment results, and feedback content. The evaluation focused on the technical correctness of the feedback, its consistency with the observed motion, clarity for users, and usefulness for practical training. Table 12 summarizes the expert evaluation results of the proposed LLM-assisted stroke feedback system. Overall, the system achieved a mean score of 8.594/10, indicating high acceptance among the evaluators. Both experts and badminton athletes consistently assigned scores above 8.30 across all evaluation criteria, suggesting that the generated feedback was perceived as technically accurate, interpretable, and practically useful from both coaching and athlete perspectives.

Although evaluation results were positive, the proposed system is not intended to replace professional coaching or elite-level performance analysis. Instead, it is an accessible decision-support tool for badminton enthusiasts, amateur players, and beginner or developmental teams. The current human-in-the-loop evaluation involved only two experts and 30 stroke videos, which may limit the generalizability of the findings. Future work will include more experts, players, and stroke samples, and incorporate quantitative evaluation metrics such as inter-rater agreement, coach approval rate, and user-perceived usefulness.

Unlike traditional rule-based systems, the proposed LLM-assisted coaching framework utilizes structured inputs consisting of stroke type, six SPIs, hierarchical fuzzy assessment results, and radar-chart ratings to generate personalized coaching feedback. Both the SPIs and the HFIS outputs are continuous-valued variables. Consequently, constructing a rule-based feedback system would require discretizing these continuous results into manually selected intervals before designing corresponding response rules and templates. Such discretization inevitably introduces additional heuristic thresholds that may influence the generated coaching responses. Furthermore, as the number of stroke types, performance indicators, and coaching knowledge increases, the number of handcrafted rules and response templates grows rapidly, making the system difficult to scale and maintain. Therefore, rather than comparing different natural-language generation approaches, this study focuses on validating whether the LLM-generated coaching feedback is technically correct, consistent with expert judgments, and practically useful through human-in-the-loop evaluation.

6. Conclusions

This study presented an explainable AI framework for badminton stroke-quality assessment using vision-based motion analysis and LLM-assisted coaching analytics. The proposed framework integrates stroke event segmentation, human pose estimation, motion feature extraction, stroke type recognition, hierarchical fuzzy assessment, and natural-language coaching feedback generation within a unified pipeline. Six SPIs were proposed to characterize badminton stroke execution from multiple biomechanical perspectives, including swing velocity, trajectory smoothness, wrist posture, elbow posture, footwork efficiency, and racket control. To improve interpretability and reduce the complexity of multi-criteria assessment, an HFIS was developed to progressively aggregate the six SPIs into three intermediate biomechanical assessments, including swing mechanics, joint coordination, and stroke control, and then generate an overall stroke-quality score. Furthermore, a knowledge-guided LLM-assisted coaching module was introduced to transform the structured assessment results into grounded, personalized, and human-readable coaching feedback. Experimental results demonstrated the effectiveness and robustness of the proposed framework. The proposed mST-GCN++ achieved reliable stroke recognition performance under both random-split and leave-one-subject-out evaluation protocols. The proposed stroke-quality assessment framework also showed strong agreement with expert evaluations, as indicated by correlation, intraclass correlation, and error analyses. In addition, a supplementary binary analysis of effective and ineffective strokes suggested that the proposed assessment scores are generally consistent with practical rally outcomes under the adopted evaluation protocol. Finally, the human-in-the-loop evaluation involving badminton experts and players confirmed that the generated coaching feedback is technically appropriate, interpretable, and practically useful for badminton training. Overall, the proposed framework demonstrates that structured biomechanical representations, transparent fuzzy reasoning, and large language models can be effectively integrated into an explainable AI system for personalized sports coaching.

Although promising results have been achieved, several limitations remain. The current framework assumes an unobstructed smartphone camera view of the player, and severe body occlusion, unfavorable camera viewpoints, poor illumination, or motion blur may degrade assessment performance. Future work will focus on enhancing the robustness and generalizability of the proposed framework through uncertainty-aware assessment under unconstrained recording conditions. In addition, the assessment framework will be extended to cover a broader range of badminton techniques, incorporate player-specific adaptive reference models, and integrate multimodal sensor information to further improve assessment accuracy, robustness, and practical coaching effectiveness.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by the National Science and Technology Council (NSTC), Taiwan, under Grant No. NSTC 112-2221-E-008-099-MY2.

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. G. Bisht, V. B. Mishra, Biomechanical performance assessment of overhead backhand smash strokes of male badminton players, *Int. J. Phys. Educ. Sports Health*, **12** (2025), 159–161. https://doi.org/10.22271/kheljournal.2025.v12.i3c.3791
2. H. Huang, Z. Guo, M. Zhao, M. Liu, J. Dai, Differences in backcourt forehand clear stroke between novice players and experienced badminton players: Based on body segment acceleration data, *BMC Sports Sci. Med. Rehabil.*, **17** (2025), 118. https://doi.org/10.1186/s13102-025-01163-w
3. X. Cui, W. K. Lam, Q. Gao, X. Wang, T. Zhao, Biomechanical analysis on skilled badminton players during take-off phase in forehand overhead strokes: A pilot study, *Biology*, **11** (2022), 1401. https://doi.org/10.3390/biology11101401
4. F. Li, S. Li, X. Zhang, G. Shan, Biomechanical insights for developing evidence-based training programs: Unveiling the kinematic secrets of the overhead forehand smash in badminton through novice-skilled player comparison, *Appl. Sci. (Basel)*, **13** (2023), 12488. https://doi.org/10.3390/app132212488

5. E. Edmizal, E. Barlian, D. Donie, A. Komaini, T. H. Sin, U. Umar, et al., Biomechanical analysis of smash stroke in badminton: A comparative study of elite and recreational players: A systematic review, *Retos*, **57** (2024), 809–817.
6. I. Ghosh, S. R. Ramamurthy, A. Chakma, N. Roy, DeCoach: Deep learning-based coaching for badminton player assessment, *Pervasive Mob. Comput.*, **83** (2022), 101608. https://doi.org/10.1016/j.pmcj.2022.101608
7. X. Zhu, L. Liu, J. Huang, G. Chen, X. Ling, Y. Chen, The analysis of motion recognition model for badminton player movements using machine learning, *Sci. Rep.*, **15** (2025), 19030. https://doi.org/10.1038/s41598-025-02771-9
8. Y. Wang, L. Wu, Design of badminton technical movement recognition system based on improved Agnes algorithm, *J. Electr. Syst.*, **20** (2024), 1981–1991. https://doi.org/10.52783/jes.3113
9. Y. Jia, N. A. Abdullah, H. Eliza, Q. Lu, D. Si, H. Guo, et al., A narrative review of deep learning applications in sports performance analysis: Current practices, challenges, and future directions, *BMC Sports Sci. Med. Rehabil.*, **17** (2025), 249. https://doi.org/10.1186/s13102-025-01294-0
10. L. Xiao, Y. Cao, Y. Gai, E. Khezri, J. Liu, M. Yang, Recognizing sports activities from video frames using deformable convolution and adaptive multiscale features, *J. Cloud Comput.*, **12** (2023), 167. https://doi.org/10.1186/s13677-023-00552-1
11. Y. H. Hsu, C. C. Yu, H. Y. Cheng, Enhancing badminton game analysis: An approach to shot refinement via a fusion of shuttlecock tracking and hit detection from monocular camera, *Sensors*, **24** (2024), 4372. https://doi.org/10.3390/s24134372
12. P. Liu, J. Wang, MonoTrack: Shuttle trajectory reconstruction from monocular badminton video, *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, New Orleans, LA, USA, (2022), 3512–3521. https://doi.org/10.1109/CVPRW56347.2022.00395
13. J. Y. Chang, BST: Badminton stroke-type transformer for skeleton-based action recognition in racket sports, preprint, arXiv:2502.21085.
14. Y. Lai, Z. Shi, C. Zhu, YO-CSA-T: A real-time badminton tracking system utilizing YOLO based on contextual and spatial attention, in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Hangzhou, China, (2025), 732–739. https://doi.org/10.1109/IROS60139.2025.11246442
15. Z. Cao, T. Liao, W. Song, Z. Chen, C. Li, Detecting the shuttlecock for a badminton robot: A YOLO based approach, *Expert Syst. Appl.*, **164** (2021), 113833. https://doi.org/10.1016/j.eswa.2020.113833
16. Y. C. Huang, I. N. Liao, C. H. Chen, T. U. İk, W. C. Peng, TrackNet: A deep learning network for tracking high-speed and tiny objects in sports applications, in *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance*, Taipei, Taiwan, (2019), 1–8. https://doi.org/10.1109/AVSS.2019.8909871
17. Y. J. Chen, Y. S. Wang, TrackNetV3: Enhancing shuttlecock tracking with augmentations and trajectory rectification, in *5th ACM International Conference on Multimedia in Asia*, Tainan, Taiwan, (2023), 1–7. https://doi.org/10.1145/3595916.3626370
18. Y. Wu, Z. Chen, H. Zhang, Y. Yang, W. Yi, Enhanced pose estimation for badminton players via improved YOLOv8-pose with efficient local attention, *Sensors (Basel)*, **25** (2025), 4446. https://doi.org/10.3390/s25144446
19. Q. Zhang, L. Cheng, Research on deep learning based badminton players' movement analysis methods, in *2024 2nd International Conference on Signal Processing and Intelligent Computing*, Guangzhou, China, (2024), 1175–1178. https://doi.org/10.1109/SPIC62469.2024.10691578

20. P. Harris, G. S. Mathi, S. Sivadarsini, B. T. Jenifer, Methods of postural assessment in badminton: A comparative analysis of conventional techniques and AI-powered pose estimation, *Int. J. Adv. Res. Comput. Sci.*, **16** (2025), 49. https://doi.org/10.26483/ijarcs.v16i3.7245
21. D. Rode, A. Dunkel, R. Willi, P. Wolf, M. Xiloyannis, R. Riener, Assessment of monocular human pose estimation models for clinical movement analysis, *Sci. Rep.*, **15** (2025), 38767. https://doi.org/10.1038/s41598-025-22626-7
22. G. H. Martínez, *Openpose: Whole-body Pose Estimation*, Ph.D thesis, Carnegie Mellon University in Pittsburgh, PA, USA, 2019.
23. J. W. Kim, J. Y. Choi, E. J. Ha, J. H. Choi, Human pose estimation using mediapipe pose and optimization method based on a humanoid model, *Appl. Sci.*, **13** (2023), 2700. https://doi.org/10.3390/app13042700
24. V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, M. Grundmann, BlazePose: On-device real-time body pose tracking, preprint, arXiv:2006.10204.
25. A. Kini, S. Rubini, PoseNet: A novel YOLO-driven framework for badminton posture detection and correction, in *International Conference on ICT for Sustainable Development*, Goa, India, (2025), 227–239. https://doi.org/10.1007/978-3-032-06662-6_22
26. W. Liu, Q. Bao, Y. Sun, T. Mei, Recent advances of monocular 2D and 3D human pose estimation: A deep learning perspective, *ACM Comput. Surv.*, **55** (2022), 1–41. https://doi.org/10.1145/3524497
27. S. Yan, Y. Xiong, D. Lin, Spatial temporal graph convolutional networks for skeleton-based action recognition, *Proc. AAAI Conf. Artif. Intell.*, **32** (2018), 7444–7452. https://doi.org/10.1609/aaai.v32i1.12328
28. L. Shi, Y. Zhang, J. Cheng, H. Lu, Skeleton-based action recognition with multi-stream adaptive graph convolutional networks, *IEEE Trans. Image Process.*, **29** (2019), 9532–9545. https://doi.org/10.1109/tip.2020.3028207
29. L. Yu, L. Tian, Q. Du, J. A. Bhutto, Multi-stream adaptive spatial-temporal attention graph convolutional network for skeleton-based action recognition, *IET Comput. Vis.*, **16** (2021), 143–158. https://doi.org/10.1049/cvi2.12075
30. U. Kilic, O. O. Karadag, G. T. Ozyer, AGMS-GCN: Attention-guided multi-scale graph convolutional networks for skeleton-based action recognition, *Knowl. Based Syst.*, **311** (2025), 113045. https://doi.org/10.1016/j.knosys.2025.113045
31. Y. Cao, C. Liu, Z. Huang, Y. Sheng, Y. Ju, Skeleton-based action recognition with temporal action graph and temporal adaptive graph convolution structure, *Multimed. Tools Appl.*, **80** (2021), 29139–29162. https://doi.org/10.1007/s11042-021-11136-z
32. Y. Chen, Z. Zhang, C. Yuan, B. Li, Y. Deng, W. Hu, Channel-wise topology refinement graph convolution for skeleton-based action recognition, preprint, arXiv:2107.12213.
33. D. Chen, M. Chen, P. Wu, M. Wu, T. Zhang, C. Li, Two-stream spatio-temporal GCN-transformer networks for skeleton-based action recognition, *Sci. Rep.*, **15** (2025), 4982. https://doi.org/10.1038/s41598-025-87752-8
34. T. Yagneshwar, S. Mukherjee, STTGC-Net: Spatial-temporal transformer with graph convolution for skeleton-based action recognition, in *14th Indian Conference on Computer Vision, Graphics and Image Processing*, New York, NY, USA, (2023), 1–10. https://doi.org/10.1145/3627631.3627640

35. M. A. Sarwar, Y. C. Lin, Y. A. Daraghmi, T. U. İk, Y. L. Li, Skeleton based keyframe detection framework for sports action analysis: Badminton smash case, *IEEE Access*, **11** (2023), 90891–90900. https://doi.org/10.1109/access.2023.3307620
36. F. Asriani, A. Azhari, W. Wahyono, Enhanced badminton stroke recognition using hybrid RGB–skeleton features and ensemble learning, *Eng. Technol. Appl. Sci. Res.*, **16** (2026), 31303–31309. https://doi.org/10.48084/etasr.15586
37. Y. Li, Y. Feng, X. Wang, G. Lu, Deep learning-based badminton action recognition and quality assessment, *Intell. Data Anal.*, **30** (2025), 531–543. https://doi.org/10.1177/1088467x251353444
38. R. Pang, X. Jiang, Research on image recognition technology to assess the quality of badminton players' stroke action and training improvement strategy, *Int. J. Comput. Inf. Syst. Ind. Manage. Appl.*, **18** (2026), 1–9. https://doi.org/10.70917/ijcisim-2026-0083
39. T. Chen, K. Chen, X. Liu, P. Ke, Z. Sun, BadminSense: Enabling fine-grained badminton strokes evaluation on single smartwatch, in *2026 CHI Conference on Human Factors in Computing Systems*, Barcelona, Spain, (2026), 1–20. https://doi.org/10.1145/3772318.3790998
40. M. Seong, G. Kim, D. Yeo, Y. Kang, H. Yang, J. DelPreto, et al., Wearable sensor-based biomechanical dataset for evaluation of badminton performance, *Sci. Data*, **11** (2024), 343. https://doi.org/10.1038/s41597-024-03144-z
41. X. Lai, Y. Lai, J. Chen, S. Huang, Q. Gao, C. Huang, Evaluation strategies for large language model-based models in exercise and health coaching: Scoping review, *J. Med. Internet Res.*, **27** (2025), 79217. https://doi.org/10.2196/79217
42. X. Lai, J. Chen, Y. Lai, S. Huang, Y. Cai, Z. Sun, et al., Using large language models to enhance exercise recommendations and physical activity in clinical and healthy populations: Scoping review, *JMIR Med. Inf.*, **13** (2025), 59309. https://doi.org/10.2196/59309
43. Z. He, J. Wang, B. Zhang, Y. Li, Knowledge-grounded large language model for personalized sports training plan generation, *Sci. Rep.*, **16** (2026), 6793. https://doi.org/10.1038/s41598-026-37075-z
44. N. Hegde, M. Vardhan, D. Nathani, E. Rosenzweig, C. Speed, A. Karthikesalingam, et al., Infusing behavior science into large language models for activity coaching, *PLoS Digit. Health*, **3** (2023), 0000431. https://doi.org/10.1371/journal.pdig.0000431
45. Y. Wang, X. Feng, Y. Wu, X. Zhang, J. Wu, CoachXNet: An artificial intelligence and internet of things integrated platform for personalized training and feedback in digital sports, *Int. J. Comput. Intell. Syst.*, **19** (2026), 49. https://doi.org/10.1007/s44196-025-01146-2
46. O. Bodemer, Enhancing individual sports training through artificial intelligence: A comprehensive review, preprint, TechRxiv:24005916.v1.
47. Y. Han, Intent-aware personalized feedback generation from coach-athlete dialogues in sports training, *J. King Saud Univ. Comput. Inf. Sci.*, **37** (2025), 150. https://doi.org/10.1007/s44443-025-00165-5
48. E. Weissblueth, A. Hirsh, I. Linder, Human expertise and artificial intelligence in coaching: A comparative analysis of feedback for triathlon front crawl technique, *Int. J. Sports Sci. Coach*, **0** (2026), 17479541261435125. https://doi.org/10.1177/17479541261435125
49. J. Li, Y. Shen, W. Zhou, H. Chen, SearchExpert: A GenAI-driven framework for reasoning-intensive multimedia information fusion through fine-tuning and reinforcement learning, *Inf. Fusion*, **126** (2026), 103665. https://doi.org/10.1016/j.inffus.2025.103665

50. X. Zhu, R. Wang, L. Cascone, E. Xiaosong, H. Chen, EUPA: Embodied uncertainty-aware multimodal affect perception for consumer electronic systems, *IEEE Trans. Consum. Electron.*, **72** (2026), 6508–6515. https://doi.org/10.1109/TCE.2026.3679325
51. X. Zhu, Z. Liu, E. Cambria, X. Yu, X. Fan, H. Chen, et al., A client–server based recognition system: Non-contact single/multiple emotional and behavioral state assessment methods, *Comput. Methods Programs Biomed.*, **260** (2025), 108564. https://doi.org/10.1016/j.cmpb.2024.108564
52. Q. Li, R. Li, K. Ji, W. Dai, Kalman filter and its application, in *2015 8th International Conference on Intelligent Networks and Intelligent Systems*, Tianjin, China, (2015), 74–77. https://doi.org/10.1109/ICINIS.2015.35
53. H. Duan, J. Wang, K. Chen, D. Li, PYSKL: Towards good practices for skeleton action recognition, in *30th ACM International Conference on Multimedia*, Lisboa, Portugal, (2022), 7351–7354. https://doi.org/10.1145/3503161.3548546
54. Y. A. Farha, J. Gall, MS-TCN: Multi-stage temporal convolutional network for action segmentation, in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, (2019), 3575–3584.
55. Y. Nie, N. H. Nguyen, P. Sinthong, J. Kalagnanam, A time series is worth 64 words: Long-term forecasting with transformers, preprint, arXiv:2211.14730.
56. H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*, 4th edition, Springer Dordrecht, New York, 2011. https://doi.org/10.1007/978-94-010-0646-0
57. R. Khanam, M. Hussain, YOLOv11: An overview of the key architectural enhancements, preprint, arXiv:2410.17725.