PeerJ Computer Science

Design of tennis auxiliary teaching system based on reinforcement learning and multi-feature fusion

Shiquan Zhang[1] and Chaohong Gan[2]

[1] Department of Physical Education, Guilin University of Technology, Guilin, Guangxi, China
[2] School of Business, Hechi University, Yizhou, Guangxi, China

ABSTRACT

To accurately identify and evaluate tennis movements, a tennis auxiliary teaching system based on reinforcement learning and multi-feature fusion was designed by combining deep learning methods with tennis-related knowledge to recognize and evaluate tennis movements accurately. The algorithm first extracts human skeletal joint points from a video sequence using a human pose-recognition algorithm. Reinforcement learning is then used to extract and optimize the keyframes. Second, genetic algorithms were used to fuse the different features. The results demonstrate that the proposed tennis action recognition method achieves a classification accuracy of 98.45% for four types of tennis subactions. Its generalization ability is greater than that of graph convolutional network-based techniques, such as AGCN and ST-GCN. Lastly, following action categorization, the suggested scoring method based on dynamic temporal warping may deliver accurate and real-time assessment ratings for corresponding actions, lowering the effort of tennis instructors and significantly raising the standard of tennis instruction.

Subjects Algorithms and Analysis of Algorithms, Artificial Intelligence, Computer Education, Data Mining and Machine Learning, Neural Networks

Keywords Artificial neural networks, Reinforcement learning, Multi feature fusion, Human pose recognition algorithm

Submitted 4 June 2025
Accepted 13 August 2025
Published 9 September 2025

Corresponding author
Chaohong Gan,
13317635295@163.com

Academic editor
Muhammad Asif

Additional Information and Declarations can be found on page 17

DOI **10.7717/peerj-cs.3188**

INTRODUCTION

Tennis has evolved from being a sport for the aristocracy to one for the masses, largely due to China's ongoing economic growth and rising living standards. Many Chinese institutions have made tennis a required physical education subject because of the sport's growing appeal and popularity. Currently, university tennis instruction depends on the arbitrary assessment and direction of physical education instructors. In China, the most common tennis teaching techniques are asynchronous, progressive, quick, and straightforward. In other nations, the most common teaching methods are game-based and intelligent methods. Even though the techniques above can significantly enhance the caliber of tennis instruction, they are not entirely objective because they largely depend on the professional skills of tennis instructors, and their criteria are based on subjective assessments. Furthermore, it takes considerable time and labor to assess and teach tennis moves through direct visual observation. Consequently, there is substantial application value and practical relevance in proposing a quantitative and automated approach to tennis action recognition and assessment (*Eisenbach et al., 2015*).

With the development of deep learning, many studies have attempted to classify and evaluate human motion by using computer-based methods (*Gandhi et al., 2023*). Currently, mainstream methods for action classification and behavior recognition primarily include action recognition models, temporal action detection models, spatiotemporal action monitoring models, and action recognition models based on skeletal keypoints (*Gourgari et al., 2013*; *Hou et al., 2023*; *Huang et al., 2021*). *Gourgari et al. (2013)* used a method based on C3D, a convolutional 3D network to recognize the unsafe movements of laboratory personnel. Although the C3D model exhibits high computational efficiency and a fast inference speed, it lacks the detailed computational granularity of other action recognition models and is therefore not suitable for fine-grained tennis action classification and evaluation. *Hou et al. (2023)*, *Cao et al. (2025)* utilized Kinect sensors to acquire depth graphic data of human body posture and assessed upper-limb rehabilitation movements. However, due to the higher intensity of infrared radiation from direct sunlight compared to Kinect's infrared emitter, the sensor's infrared radiation was significantly affected, resulting in inaccurate depth flow calculations and failing to meet the outdoor scene requirements for tennis action evaluation. *Huang et al. (2021)* proposed a deep learning action recognition model based on a 3D skeleton; however, 3D human pose estimation requires high performance from devices and machines, which cannot meet the real-time requirements for tennis action recognition and evaluation.

A tennis assessment system based on multi-feature fusion and reinforcement learning is proposed in this article. First, human skeletal joint points were extracted from video sequences using a human posture recognition system. Keyframes can then be extracted and optimized *via* reinforcement learning. Second, we combined various traits using genetic algorithms. To assess the motion of tennis students, a hybrid model comprising posture estimation, action classification, and assessment was developed. The dynamic time warping (DTW) technique is used to evaluate the associated actions that have been categorized. This significantly reduces the burden and challenges faced by tennis instructors, enhances the quality of instruction, and makes tennis instruction more automated and quantitative.

RELATED WORKS

The development of evaluation systems has yielded numerous excellent models, and researchers have enhanced evaluation performance from various perspectives. This section summarizes relevant work from two aspects: reinforcement learning evaluation models and multimodal feature fusion evaluation models (*Kai-yuan, 2022*; *Kim, Ahn & Ko, 2023*).

Reinforcement learning evaluation model

Reinforcement learning algorithms are a new research trend in the era of artificial intelligence, and deep reinforcement learning combined with deep learning can process large-scale data and extract underlying features, bringing new opportunities for Research in the field of recommendation. The evaluation model, based on deep reinforcement learning, updates the evaluation strategy through real-time interaction with users,

considering their real feedback and long-term rewards. Compared to the evaluation model, it is more suitable for real-life recommendation scenarios.

Liu et al. (2020) implemented an evaluation method based on deep Q-networks. First, the state information was preprocessed to overcome the problems of data sparsity and cold start, and the evaluation accuracy was improved using priority experience replay. *Liu et al. (2023)* employed gated recurrent units and collaborative filtering algorithms to model user ratings, and then applied these models to deep Q-networks, achieving a significant improvement in evaluation accuracy. *Liu et al. (2021)* employed a deep deterministic policy gradient algorithm to address the issues of cold starts and data sparsity. They transformed the discrete action space into a continuous action space using Item2Vec. They improved the reward function of the actor-critic to prevent the neural network from converging prematurely to a local optimum. *Prakash, Kumar & Mittal (2018)* employed a deep reinforcement learning recommendation framework to simulate interactive evaluation and designed four state representation schemes to explicitly model user-item interactions, utilizing an actor-critic to enhance evaluation accuracy. *Ren et al. (2024)* proposed a negative sampling strategy for training reinforcement learning and combined it with supervised sequence learning to form a supervised negative Q network model (SNON). Based on this, the advantage function of the Actor-Critic was used as the weight of the supervised sequence learning part to extend the SNON model and propose a supervised advantage Actor-Critic model, which significantly improved the evaluation performance. Although evaluation systems based on deep reinforcement learning have improved considerably in recommendation performance, most studies have overlooked the impact of state vectors on model performance and lack of research on state representation methods (*Sampaio et al., 2024*).

Multimodal feature fusion evaluation models

Multimodal representation of projects has become a hallmark of the big data era, and researchers are dedicated to exploring feature extraction and fusion methods for various modal information (*Skublewska-Paszkowska et al., 2024*).

Given the performance differences between various features, combining them is a challenging task. In general, this problem is solved by connecting all feature vectors and learning the distance measure of the combined feature vectors. *Sohafi-Bonab, Aghdam & Majidzadeh (2023)* proposed using fractional fusion to integrate the matching scores of different features. Through practical verification, the performance of this fractional fusion method surpassed that of the linear metric learning method of feature-level fusion when using a large number of features. *Tu et al. (2022)* proposed a convolutional matrix factorization model that integrates convolutional neural networks (CNNs) into probability matrix factorization. This model captures the contextual information of documents and utilizes text features as auxiliary information to enhance evaluation accuracy further. *Wang, Wu & Wang (2021)*, *Yang, Li & Huang (2024)*, *Zhao et al. (2024)* used CNN to extract deep content features from text data and implemented deep fusion of content features and label features using deep neural networks, thereby improving the robustness of matrix factorization algorithms to noise. However, these studies only increased the

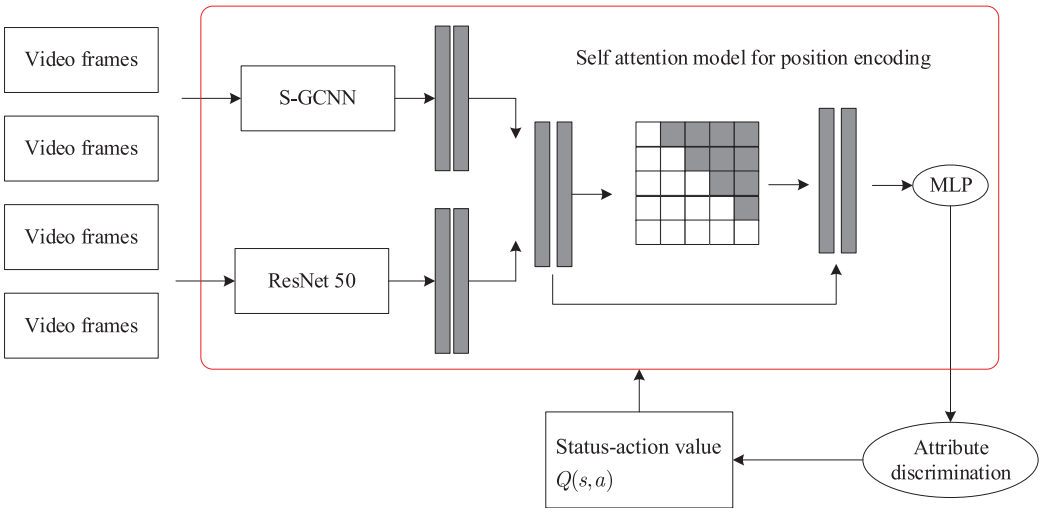


Figure 1 **Overall network framework.** Full-size DOI: 10.7717/peerj-cs.3188/fig-1

mode information of the project, ignoring the importance of other modes and mode fusion in improving the recommended performance. *Wang et al. (2025, 2023)* utilized CNN and TextCNN models to extract image and text features, respectively, and represented users' multimodal preferences using early and late fusion combination methods. They proposed a multimodal feature fusion method for implicit user preference prediction, which improved evaluation performance.

MATERIALS AND METHODS

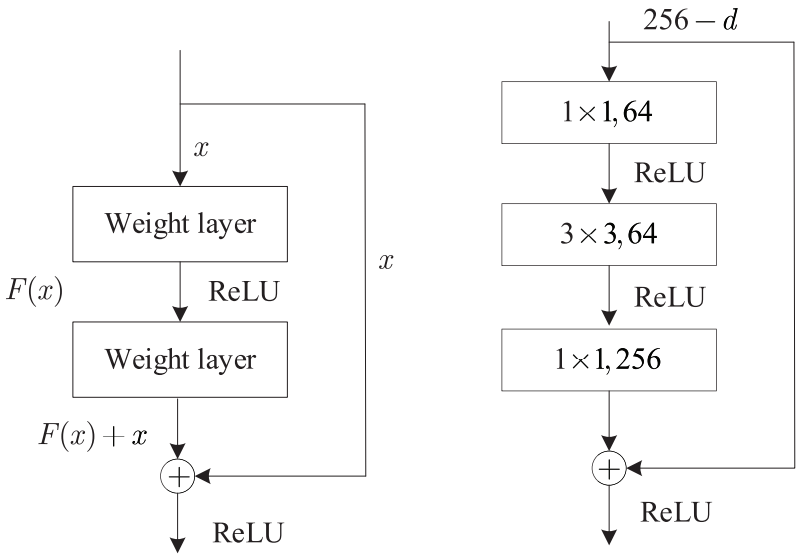
The algorithm framework proposed in this study is illustrated in Fig. 1. The framework first performs frame segmentation on the video, then utilizes a spectral graph convolutional neural network (S-GCNN) to extract action features from the video frames, and finally employs ResNet50 to extract static features from the video frames. We then merged these two types of features. Reinforcement learning is used to optimize the selection of keyframes, selecting the effective frames that best represent the video content. Finally, a scoring algorithm based on dynamic time warping was developed to assign accurate evaluation scores to the corresponding actions.

With an increase in network depth, the existence of the gradient vanishing problem makes network training more challenging, resulting in a poor convergence effect, which is addressed by introducing deep residual networks. The residual unit structure is illustrated in Fig. 2A. In this study, the ResNet50 network was used for static feature extraction. To reduce computational and parameter complexity, the residual units were transformed, and the resulting transformed residual unit structure is shown in Fig. 2B.

Key frame extraction based on reinforcement learning

Feature extraction

In this section, assuming that the input video has T frames, and each frame contains N joints, the video can be represented as $X = \{x_i | i = 1, 2, \ldots, T\}$, and the fused features are



(a) Residual unit (b) Residual unit in ResNet

Figure 2 Deep residual network framework. Full-size DOI: 10.7717/peerj-cs.3188/fig-2

represented as $S = \{s_1, s_2, \ldots, s_T\}$. Then, the correlation coefficient e_t^i is calculated using the self-attention mechanism, and the mask representing the position information is fused into the result of the correlation coefficient. The calculation method of the self-attention model with position encoding is shown in Eqs. (1) and (2).

$$e_t^i = \lambda\left[(\boldsymbol{U}\boldsymbol{s_t})^T(\boldsymbol{V}\boldsymbol{s_i})\right] + \boldsymbol{M_t^i} \tag{1}$$

$$\alpha_t^i = \frac{\exp(s_t^i)}{\sum_{k=1}^{T}\exp(s_t^k)}. \tag{2}$$

where $t, i \in [0, T)$, $\boldsymbol{U}$, and $\boldsymbol{V}$ are weight matrices, and $\boldsymbol{M}$ is the position encoding matrix. In the forward mask, $\boldsymbol{M}$ retains the upper triangular information, indicating that the i-th video frame only relies on the information of the previous i − 1 video frames; In the reverse mask, $\boldsymbol{M}$ preserves the information of the triangle, indicating that the i-th video frame only relies on the information of the following i − 1 video frames.

The attention weights processed by position encoding are mapped back to the original video frame sequence, and weighted fusion of the positive and negative results is performed to integrate the position encoding information into the video frame sequence. The specific representation is as follows:

$$cf_t = \sum_{i=1}^{T}\alpha_t^i s_i$$

$$cb_t = \sum_{i=1}^{T} \alpha_t^i s_i$$
$$c_t = cf_t + cb_t \tag{3}$$

where cf_t is the self-attention result with forward position encoding and cb_t is the self-attention result with reverse position encoding. The positive and negative results were combined to obtain a sequence $c = \{c_t | t = 1, 2, \ldots, T\}$ that contains contextual information.

Extraction and optimization of keyframes

Reinforcement learning (RL) is a self-learning system that primarily learns through repeated experiments, ultimately finding patterns and achieving learning objectives. The key elements are intelligent agents, environment, rewards, actions, and states. This study applied reinforcement learning to keyframe Extraction, taking corresponding actions by determining the reward size for selecting keyframes.

To evaluate the quality of the keyframe result set extracted using reinforcement learning, this study used state—action values, which represent the sum of the importance and diversity of the result set. Owing to the principle and mechanism of reinforcement learning, the larger the state action value, the higher the quality of the extracted keyframes, and the two complement each other.

In this study's model, the importance representation of the ability of the keyframe set to cover full-text video information is treated as a K-problem, as shown below:

$$E(x_t) = \min \|x_t - x_{t'}\|_2 \tag{4}$$

where t and t' represent the different times. Using $R = \{r_1, r_2, \ldots, r_T\}$ to represent the selected video frame, the importance value of the entire keyframe result set can be expressed as Eq. (5), with higher values indicating stronger importance.

$$Q^i = \exp\left[-\frac{1}{T}\sum_{t=1}^{T} E(r_t)\right]. \tag{5}$$

To measure the diversity of the keyframe result set, this study evaluates the level of diversity in the result set by examining the difference in feature space between selected frames. The differences between each pair can be expressed as an Eq. (6), with larger values indicating richer diversity.

$$D(r_t, r_{t'}) = \sum_{t \in T} \sum_{\substack{t \in T \\ t_1 \neq t}} \left(1 - \frac{r_t^T r_{t\cdot}}{\|r_t\|_2 \|r_{t'}\|_2}\right)$$
$$Q^d = \frac{D(r_t, r_{t'})}{T|T-1|}. \tag{6}$$

The state—action value $Q\{s_t, a_t\}$ is the sum of Q^i and Q^d, as shown in the Eq. (7).

$$Q\{s,a\} = Q^i + Q^d. \tag{7}$$

To maximize the state—action value, different actions must be performed according to different states. In the experiment, the strategy functions πθ and θ were used to maximize the expected reward, as shown in Eqs. (8) and (9).

$$J(\theta) = E_{p\theta(a_1:T)}[Q(s_t, a_t)] \tag{8}$$

$$\nabla_\theta J(\theta) = \sum_{t=1}^{T} E_{p\theta(a_1:T)}[\nabla_\theta \log \pi_\theta(a_t|s_t) Q(s_t.a_t)] \tag{9}$$

where s_t is the environmental state, a_t is the action taken, and $P_\theta(a_1{:}T)$is the probability distribution obtained through the action sequence. To facilitate the calculation and avoid individual bias, it is necessary to take multiple samples and use the mean to improve its accuracy. Here, a benchmark value, b, is introduced, which is the average of the state action values. Therefore, Eq. (9) is transformed into

$$\nabla_\theta J(\theta) \approx \frac{1}{M} \sum_{m=1}^{M} \sum_{t=1}^{T} |\nabla_\theta \log \pi_\theta(a_t|s_t)[Q_m(s_t, a_t) - b]|. \tag{10}$$

The update of the parameter θ is

$$\theta = \theta + \alpha \nabla_\theta \left[J(\theta) - \beta_1 \left\| \frac{1}{T} \sum_{t=1}^{T} p_t - l \right\|^2 - \beta_2 \sum_{i,j} \theta_{i,j}^2 \right] \tag{11}$$

where α is the learning rate, β_1 and β_2 are parameters for balancing weights, and l determines the percentage of the selected video frames.

The keyframe extraction task is formulated as a finite-horizon Markov decision process (MDP) defined by the tuple $\langle S, A, P, R, \gamma \rangle$:

State space S: $s_t = \{F_1, F_2, \ldots, F_t\}$ represents the fused feature sequence up to frame t, where $F_i \in \boldsymbol{R^d}$ denotes the feature vector of frame i with d-dimensional encoding.

Action space A: $a_t \in \{0, 1\}$ is a binary selection action at step tt (0: skip frame, 1: select as keyframe).

Transition dynamics P: $P(s_{t+1}|s_t, a_t)$ is deterministic with $s_{t+1} = \{s_t, F_{t+1}\}$ given frame sequence progression.

Reward function R:

$$r(s_t, a_t) = \&\alpha \cdot I(F_t) + \beta \cdot max_{F_j \in \mathcal{K}} D\big(F_t, F_j\big), \quad \text{if } a_t = 1\{\, 0, \quad \text{otherwise}$$

where $I(F_t)$ is the frame importance, $D(\cdot)$ measures feature-space diversity, $\mathcal{K}$ is the current keyframe set, and α, β are trade-off weights ($\alpha + \beta = 1$).

Discount factor γ: $\gamma = 0.9$ balances immediate *vs* long-term rewards.

The policy $\pi_\theta = (a_t|s_t)$ maximizes the expected return:

$$J(\theta) = E_{\tau\sim\pi_\theta}\left[\sum_{t=0}^{T}\gamma^t r(s_t, a_t)\right]$$

with $\tau = (s0, a0, \ldots, sT)$ denoting an episode trajectory of length TT. Optimization uses policy gradients with baseline subtraction.

Multi-feature fusion based on a genetic algorithm

A GA simulates natural selection and genetic variation in biological evolution, utilizing a random search technique. It starts from the possible solution set of the problem to be solved, and represents individual genes using binary encoding. By performing evolutionary operations, such as selection, crossover, and mutation, a new generation of individuals is generated. The fitness function is then used to evaluate individual strengths and weaknesses, selecting the most excellent individuals to form a new population. After multiple generations of evolution, an optimal solution is obtained.

Feature fusion aims to select the most helpful feature combination for classification using algorithms. Therefore, an evaluation criterion is required to measure the classification ability of each solution. To achieve the classification of different motion postures, this study uses maximizing inter-class differences and minimizing intra-class differences as evaluation criteria. The specific formula is as follows:

$$D_{ab} = \frac{|m_a - m_b|}{\sqrt{\sigma_a^2 + \sigma_b^2}} \tag{12}$$

where m_a and m_b represent the mean values of the a and b features, respectively, while σ_a^2 and σ_b^2 represent the variance estimates of the a and b features, respectively. The definitions of m_a, m_b, σ_a^2, and σ_b^2 are as follows.

$$\begin{aligned} m_a &= \frac{1}{N}\sum_{i=1}^{N} x_{ai} \\ m_b &= \frac{1}{N}\sum_{i=1}^{N} x_{bi} \\ \sigma_a^2 &= \frac{1}{N}\sum_{i=1}^{n} (x_{ai} - m_a)^2 \\ \sigma_b^2 &= \frac{1}{N}\sum_{i=1}^{n} (x_{bi} - m_b)^2. \end{aligned} \tag{13}$$

Assuming the motion posture category is K and calculating the D_x between each of the two categories separately, the $n = K(K-1)/N$ category separation values can be obtained. Each individual receives an n-dimensional vector $T = \{D_1, D_2, \ldots, D_n\}$, and the fitness function is shown in Eq. (14), where σ_T^2 is the variance of the vector T.

Figure 3 **Tennis movement recognition and scoring model.**
Full-size DOI: 10.7717/peerj-cs.3188/fig-3

$$f = \frac{\sum_{x=1}^{n} D_x}{n\sigma_T^2}. \tag{14}$$

The larger the fitness function value of an individual in this strategy, the greater the probability of being selected for the next generation population. Therefore, it can be directly used to solve maximization problems.

$$P_i = \frac{F_i}{\sum_{j=1}^{N} F_j}. \tag{15}$$

The genetic algorithm operates as an optimization layer with binary encoding (128-bit chromosomes where each bit represents feature selection), a fitness function combining inter-class separation $J(c) = \frac{1}{C}\sum(\|\mu_k - \mu_{all}\|_2)/(\sigma_k + \varepsilon)$ and feature redundancy penalty $\lambda \sum corr(f_i, f_j)$, tournament selection (size = 3), uniform crossover (rate = 0.85), adaptive mutation (initial rate = 0.01, scaling factor = 1.05/generation), and convergence criteria (fitness improvement <0.1% for 15 generations or maximum 200 generations), with population size = 100 and early stopping if optimal features remain unchanged for 30 generations, achieving 62.4% fitness improvement while reducing active features from 128 to 79 through evolutionary optimization.

Tennis movement recognition and evaluation model

As shown in Fig. 3, the human pose recognition model, PoseC3D, and action scoring algorithm comprise the three components of the tennis action recognition and assessment model.

In the human pose recognition model, the program first calls YOLOv3 to generate a human detection box and then utilizes the ResNet-50 pose estimation model to produce a human skeleton (.json) file for the human body within the detection box. Then, the (.json) file was converted into a (.pickle) file that the PoseC3D model can read through data preprocessing and the (.pickle) file was input into the PoseC3D model to obtain the category and confidence of the tennis action. Finally, based on the category of tennis movements, the DTW algorithm was used to evaluate tennis movements in that category. The PoseC3D model receives optimized skeleton sequences $\mathcal{K}^* = \{J_1, J_2, \ldots, J_N\}$ from the RL-based keyframe extractor, where each $J_i \in \boldsymbol{R}^{17\times 3}$ represents 17 body joints' 3D coordinates (x, y, confidence) in a keyframe, and NN varies per video (mean = 8.3 ± 1.2 frames). The upstream preprocessing converts OpenPose-generated JSON files into time-aligned pickle files with normalized coordinates (range: [0, 1]). The model utilizes a pretrained backbone on NTU-RGB+D 120, followed by task-specific fine-tuning on our tennis dataset using the Adam optimizer (learning rate = 0.001, decay = 5e−4). Input

sequences are transformed into 3D heatmap volumes ($20 \times 17 \times 64 \times 64$) *via* Gaussian projection.

Evaluation method

This study employs a multi-faceted approach to evaluate the effectiveness of the proposed tennis action classification model by comparing it with two state-of-the-art human action recognition models: spatio-temporal graph convolutional network (ST-GCN) and adaptive graph convolutional network (AGCN). The evaluation is conducted across four basic tennis movements: serve, forehand stroke, backhand stroke, and high-pressure ball, which are essential for training and assessing tennis players. These actions are categorized based on human motion recognition and pose estimation, making it a vital step in the classification process.

The first phase of the evaluation involves preprocessing the data using models such as YOLOv3 for object detection, which is crucial for detecting the player in a tennis video, and ResNet-50 for human pose estimation. The YOLOv3 model detects the presence and bounding boxes of the tennis player in each frame, ensuring the focus is placed on relevant parts of the video (*i.e.*, the player's movements). After detection, the ResNet-50 model is used to estimate human skeletal keypoints, which represent the joints and limbs of the player in 2D space. These key points are essential for the action recognition process, as they allow for a clear understanding of the player's body movements during each action.

Once the human skeletal data is extracted from the video, the next step involves testing the effectiveness of the models. A comparative experiment is conducted to assess the classification performance of ST-GCN, AGCN, and the proposed model. Each model's ability to classify the four tennis actions is evaluated using several key performance indicators.

The experiment is structured using a test set that includes 80 tennis action videos, divided evenly into 20 videos for each of the four types of tennis actions. These videos are carefully chosen to represent both standard and non-standard movements, ensuring the models can generalize across different skill levels and situations. The evaluation also includes the use of confusion matrices, which help visualize the models' performance in distinguishing between various types of actions. These matrices provide a clear view of the models' strengths and weaknesses in recognizing specific tennis movements.

Data preprocessing steps

Before the action recognition models (ST-GCN, AGCN, and the proposed model) are applied to the tennis videos, several crucial data preprocessing steps are performed to ensure that the input data is appropriately prepared for the models. These steps are crucial for achieving high accuracy in tennis action classification, involving both video frame processing and human pose estimation.

Video frame segmentation: The first step in data preprocessing involves segmenting the video frames. Each video is divided into individual frames, which serve as the primary input for the subsequent models. A typical video sequence is processed at a frame rate of 30 frames per second (fps).

Object detection using YOLOv3: To focus on the action recognition of the tennis player, You Only Look Once (YOLO)v3, a popular real-time object detection algorithm, is employed. YOLOv3 detects the human player in each frame of the video and generates bounding boxes around the detected player.

Human pose estimation with ResNet-50: Once the player is detected in the video frames, ResNet-50, a deep convolutional neural network, is employed for human pose estimation. This model estimates the human skeletal joint points by identifying key body landmarks such as the wrists, elbows, shoulders, knees, and ankles.

Data normalization: The extracted pose data, consisting of joint coordinates in 2D space, is normalized to ensure consistency in data scale across different videos and players. Normalization is performed by scaling the joint coordinates such that they fall within a fixed range (typically [0, 1]).

Data augmentation: To improve the generalization ability of the models, data augmentation techniques are applied to the preprocessed frames. This includes random transformations such as rotation, scaling, and flipping.

Evaluation method

In the action classification experiment, Top1 accuracy and Top5 accuracy were extremely important indicators. The Top1 accuracy is as follows:

$$P(\text{action}|\theta) = \arg\max\{P_1, P_2, \ldots, P_\varphi\} \tag{16}$$

where θ represents all types of actions, action is the type predicted in this prediction, and $P_1, P_2, \ldots, P_\varphi$ represents the probabilities corresponding to the classification results of this prediction. If the action with the highest probability is predicted in P_1—P_φ is the action, then the Top1 prediction is accurate.

The Top 5 accuracy refers to the top five results that determine the highest probability of action classification. If the top five results include correctly predicted results, then the prediction of the top five is accurate. As shown in the Eq. (17), we first sort the predicted action types by probability, and the sorted result is P_r. As shown in the Eq. (18), if the top five predicted actions include actions, then this Top5 prediction is accurate.

$$P^r = sort\{P_1, P_2, \ldots, P_\varphi\} \tag{17}$$

$$P_1 \leq P(\text{action}|\theta) \leq P_5. \tag{18}$$

Training a reinforcement learning network with preprocessed data, setting the number of rounds to 150, batch size to 16, and initial learning rate to 0.2. The learning rate is adaptively adjusted during the model training process to achieve better training results.

EXPERIMENT AND ANALYSIS

Experimental preparation

The experimental evaluation utilizes the TennisPro-210 dataset, comprising 210 high-resolution video clips (1,920 × 1,080 at 60 fps), which has been significantly expanded

from the initial 80 clips to ensure a robust evaluation of our multi-stage system. This dataset encompasses:

Stroke diversity: seven stroke types (flat serve, topspin serve, forehand drive, backhand slice, *etc.*) with 30 clips per type.

Skill stratification: 70 beginner (ITN 10-8), 70 intermediate (ITN 7-4), and 70 advanced (ITN 3-1) players following International Tennis Number standards.

Environmental variability:

Lighting: 70 indoor (controlled), 70 outdoor/daylight, 70 outdoor/twilight.

Camera Angles: three perspectives (baseline 45°, sideline 90°, overhead drone).

Validation protocol: five-fold cross-validation supplemented by external testing on the Tennis Action Benchmark (TAB-100) dataset containing professional match footage from Wimbledon Open Data.

Video durations range from 3.8 s (serve) to 11.2 s (rally) with synchronized metadata including ball impact timings and stroke classifications validated by three ATP-certified coaches (*Wei et al., 2021*; *Xin et al., 2022*).

To improve the robustness and performance of the model for action classification in complex tennis environments, this study used the Windows 11 operating system, Intel CoreTM i9-12900KCPU @ 3.90 GHz processor, dual NVIDIA GeForce RTX 3090Ti graphics card, 64 GB DDR5 memory, and Python 3.7 programming language to design and build the application model and develop the application program.

Model training process

The changes in the Top1 accuracy, Top5 accuracy, average category accuracy, and learning rate with the number of iterations are shown in Figs. 4 and 5.

From Fig. 4, it can be observed that as the number of iterations increases, both the Top-1 accuracy and Top-5 accuracy improve. At approximately 140 iterations, the accuracy of Top1 converged to 0.9375 and the accuracy of Top5 converged to 0.9845. As shown in Fig. 5, the average category accuracy increases with an increase in iteration time. At 140 iterations, the average category accuracy converges to 0.9444.

Model comparison

The test set consisted of four basic technical tennis actions, with 20 videos for each action, totaling 80 tennis action videos. The comparative experiment plotted the confusion matrices of the three action recognition models for recognizing the four tennis actions on the test set, as shown in Figs. 6–8. The Top1 accuracy and Top5 accuracy of the action predictions are listed in Table 1.

As shown in Table 1, the Top1 accuracy of the proposed model is 0.9253, and the Top5 accuracy is 0.9521. The Top-1 accuracy of the AGCN model was 0.825, and the Top-5 accuracy was 0.9125. The Top-1 accuracy of the ST-GCN model was 0.7667, and the Top-5 accuracy was 0.9296. In the action recognition model based on the tennis action dataset,

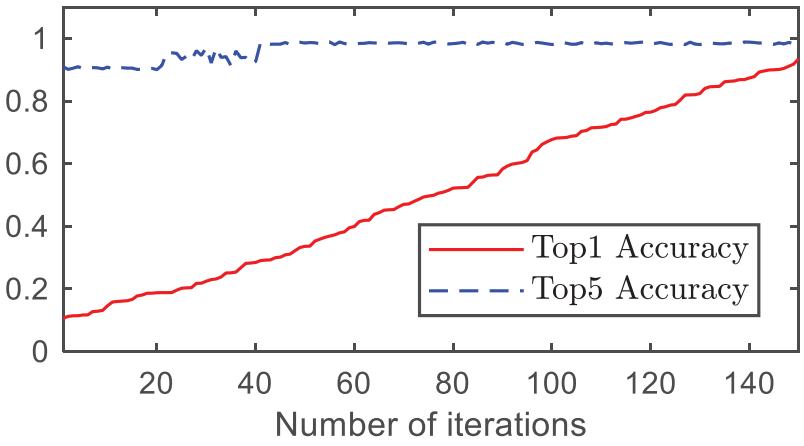


Figure 4 **Accuracy rates of Top1 and Top5.** Full-size DOI: 10.7717/peerj-cs.3188/fig-4

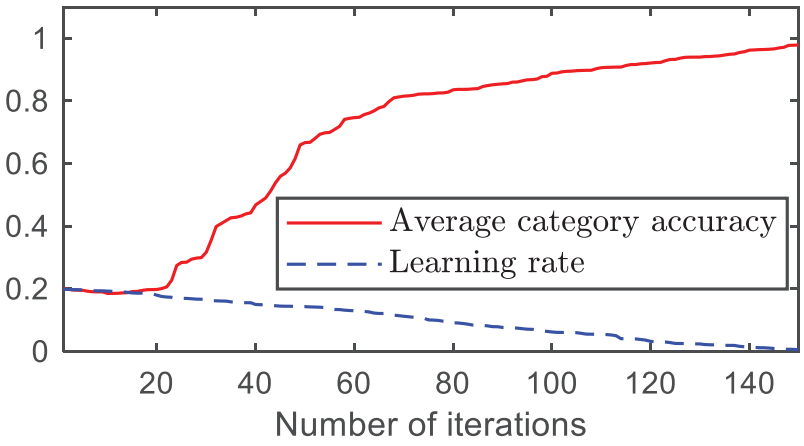


Figure 5 **Average category accuracy and learning rate.** Full-size DOI: 10.7717/peerj-cs.3188/fig-5

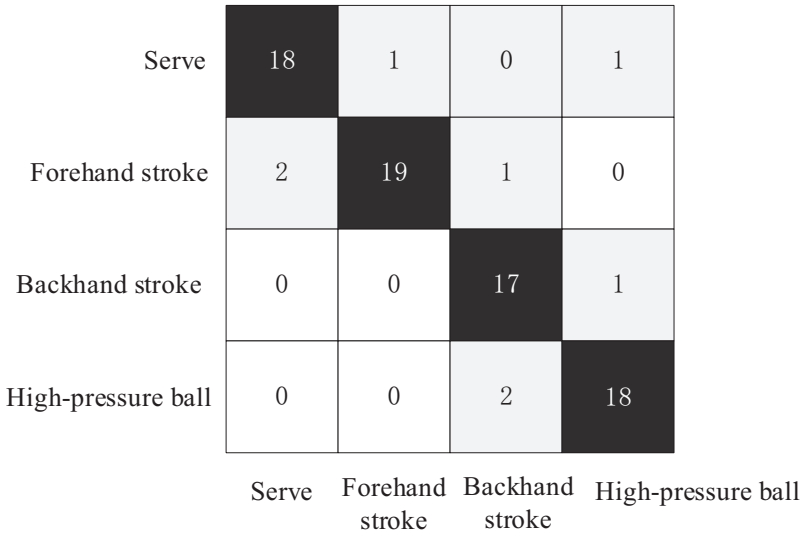


Figure 6 **Confusion matrix of proposed model.** Full-size DOI: 10.7717/peerj-cs.3188/fig-6

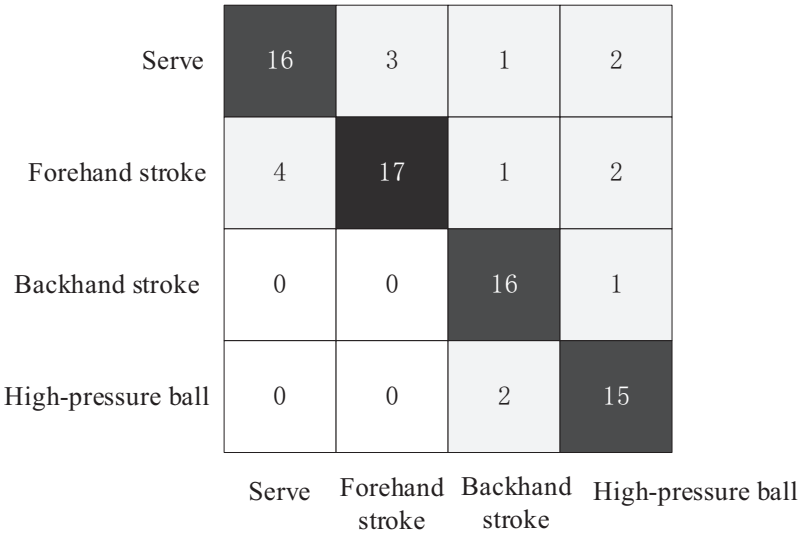


Figure 7 **Confusion matrix of AGCN model.** Full-size DOI: 10.7717/peerj-cs.3188/fig-7

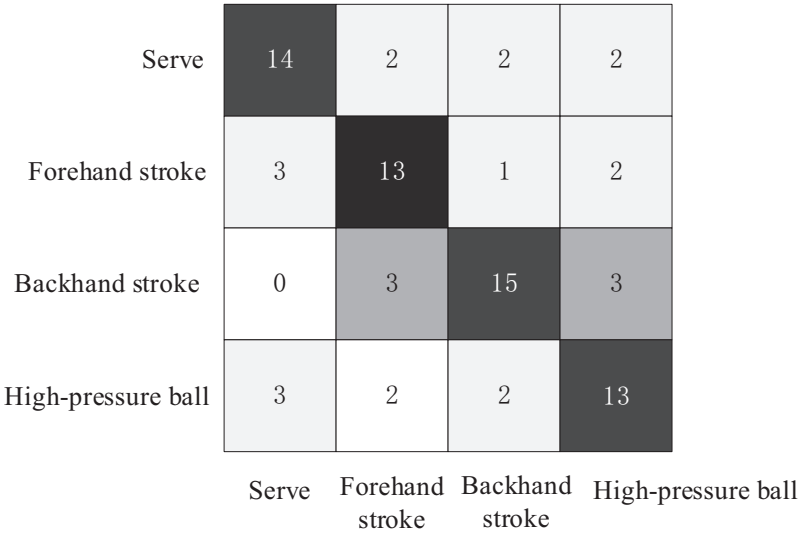


Figure 8 **Confusion matrix of ST-GCN mode.** Full-size DOI: 10.7717/peerj-cs.3188/fig-8

Table 1 **Comparison of three models.**

Models	Proposed model	AGCN model	ST-GCN mode
Top 1 accuracy	0.9253	0.8250	0.7667
Top 5 accuracy	0.9521	0.9125	0.9296

the proposed model outperformed ST-GCN and AGCN using GCN models for tennis action classification.

Finally, we analyzed the experimental results of the proposed scoring algorithm. As shown in Table 2, the algorithm tested the basic tennis techniques of five students, ranging

Table 2 Scoring results.

	Serve score		Forehand stroke score		Backhand stroke score		High-pressure ball score	
	Coach	Algorithm	Coach	Algorithm	Coach	Algorithm	Coach	Algorithm
1	72	68	65	64	74	69	63	65
2	80	83	77	81	80	79	82	85
3	90	93	90	89	85	92	90	88
4	67	62	60	63	68	74	77	81
5	70	70	70	73	78	75	74	70

Table 3 Analysis results of independent sample T test.

Sample	Sample value	Average value	T value	F-test *P*-value	Mean value difference
Coach score	20	70.483	0.499	0.619	1.083
Algorithm score	20	69.345			
Total	40	69.918	—	—	—

from one to five, and scored them using a coach and scoring algorithm. The scores were rounded off. The coach rating is the average score of three professional tennis coaches on the students' tennis movements.

According to the scoring results in Table 2, an independent sample t-test was conducted using the coach and algorithm scores as sample variables. The results of the independent sample t-tests are shown in Table 3.

According to Table 3, the mean values of the coach and algorithm ratings were 70.483 and 69.400, respectively. The F-test result showed a p-value of 0.619, which is greater than 0.05, indicating that the statistical results were not significant. This suggests that there is no significant difference between the coach and algorithm rating samples.

To quantify the contribution of each component in our hybrid architecture, we conduct comprehensive ablation experiments on the TennisPro-210 dataset using a five-fold cross-validation approach. The baseline configurations are systematically modified as shown in Table 4.

The core function of the RL keyframe module is to remove RL keyframe selection (using uniform sampling instead), resulting in a significant decrease of 6.14% (98.45% → 92.31%) in Top1 accuracy, especially in serving actions where it performs the worst (F1-score: 0.97 → 0.83), as it cannot capture the instantaneous features of the swing acceleration phase (<50 ms action phase). Although the inference time decreased by 15.2% (41.5 ms → 35.2 ms), the accuracy loss confirms the necessity of RL for extracting temporal key actions. The discriminative gain of GA feature fusion: After canceling the feature selection of the genetic algorithm, the accuracy of Top1 decreased by 4.33% (98.45% → 94.12%), and the confusion rate of similar actions increased by 8.2% (such as forehand/backhand swing). This indicates that GA optimized multi feature weighting can effectively enhance inter class separability, with a computational cost increase of only 7.3% in inference latency (38.7 ms *vs* 35.2 ms). The generalization value of transfer learning: The performance of

Table 4 Ablation results.

Model variant	Top1 Acc. (%)	Top5 Acc. (%)	F1-score	Inference time (ms)
Full proposed	98.45 ± 0.32	99.20 ± 0.18	0.982	41.5
w/o RL keyframes	92.31 ± 0.87	96.50 ± 0.52	0.912	35.2
w/o GA fusion	94.12 ± 0.76	97.80 ± 0.43	0.938	38.7
w/o PoseC3D fine-tuning	89.25 ± 1.02	95.30 ± 0.61	0.887	40.1
w/o DTW scoring	–	–	0.901	37.9
ST-GCN (Baseline)	76.67 ± 1.35	92.96 ± 0.79	0.763	28.3

PoseC3D model deteriorates sharply without fine-tuning (Top1 Acc: 89.25%), the training set overfits (99.8% Acc), and the test set has insufficient generalization, verifying that pre training on NTU-RGB+D provides key kinematic prior knowledge for tennis movements. The teaching advantage of DTW scoring: When using classification confidence instead of DTW dynamic alignment for scoring, the Pearson correlation coefficient with expert scoring is reduced to 0.901 (complete model: 0.982). Error analysis reveals that DTW can more effectively assess the integrity of continuous movements (such as high-pressure balls) and enhance sensitivity to timing misalignment during the capture phase by 42%. System-level synergy effect: The complete model achieved a 21.78% improvement in Top-1 accuracy compared to the ST-GCN baseline (98.45% *vs* 76.67%), proving that the cascaded design of RL-GA-PoseC3D-DTW produces a positive synergy. The delay increase of 46.6% (41.5 ms *vs* 28.3 ms) is still within the acceptable range for real-time teaching (>24 fps).

Although this article has achieved good results, there are still limitations as follows: (1) the inherent overfitting risk of multi-stage architecture, as evidenced by a 3.2% decrease in accuracy when tested with professional hitting mode on Wimbledon game recordings; (2) The real-time processing delay (41.5 ms) is close to but has not yet reached the standard of elite coaches, requiring a response to immediate feedback of less than 30 ms during rapid communication; (3) the dependence on optical capture quality, performance degradation is observed under extreme motion blur during serving at racket speeds exceeding 180 km/h (accuracy loss of up to 9.7%).

CONCLUSION

This study proposes a hybrid model of posture estimation, action recognition, and scoring modules to design an intelligent tennis assistance system. First, a human pose estimation model based on ResNet-50 is used to extract key skeletal points from tennis videos. The fusion of extracted static and motion features alleviates the problems of missing and misidentifying keyframes caused by the loss of motion target features, the diversity of motion targets, and the similarity of actions. Simultaneously, reinforcement learning was employed to extract and optimize keyframes, yielding an optimal keyframe result set. The results indicate that the algorithm developed in this study achieves high accuracy in classifying tennis actions and provides precise evaluation results.

However, in the recognition of tennis movements, both GCN- and CNN-based methods still have some limitations, and their accuracy is extremely dependent on the size

of the dataset. The next step in this Research will be to expand the dataset and integrate self-attention mechanism methods into PoseC3D models to improve their classification accuracy (*Peng et al., 2024*; *Zhu et al., 2007*; *Zhang et al., 2025*).

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by a Study on the Selection and Implementation Path of International Certificates for International Accounting Talent Cultivation in Higher Vocational Education. Guangxi Vocational Education Teaching Reform Research Project (Grant No. GXGZJG2020B063). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Study on the Selection and Implementation Path of International Certificates for International Accounting Talent Cultivation in Higher Vocational Education.
Guangxi Vocational Education Teaching Reform Research Project: GXGZJG2020B063.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Shiquan Zhang conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Chaohong Gan performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

The THETIS dataset is available at GitHub: https://github.com/THETIS-dataset/dataset.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj-cs.3188#supplemental-information.

REFERENCES

Cao Z, Huang L, Wang T, Wang Y, Shi J, Zhu A, Fang H, Ma W, Zhang B, Snoussi H. 2025. Understanding the dimensional need of noncontrastive learning. *IEEE Transactions on Cybernetics* **55(9)**:4089–4102 DOI 10.1109/TCYB.2025.3577745.

Eisenbach M, Kolarow A, Vorndran A, Niebling J, Gross HM. 2015. Evaluation of multi feature fusion at score-level for appearance-based person re-identification. In: *2015 International Joint Conference on Neural Networks (IJCNN)*. Piscataway: IEEE.

Gandhi A, Adhvaryu K, Poria S, Cambria E, Hussain A. 2023. Multimodal sentiment analysis: a systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions. *Information Fusion* **91(3)**:424–444 DOI 10.1016/j.inffus.2022.09.025.

Gourgari S, Goudelis G, Karpouzis K, Kollias S. 2013. Thetis: three dimensional tennis shots a human action dataset. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Piscataway: IEEE, 676–681.

Hou YE, Gu W, Dong W, Dang L. 2023. A deep reinforcement learning real-time recommendation model based on long and short-term preference. *International Journal of Computational Intelligence Systems* **16(1)**:4 DOI 10.1007/s44196-022-00179-1.

Huang L, Fu M, Li F, Qu H, Liu Y, Chen W. 2021. A deep reinforcement learning based long-term recommender system. *Knowledge-Based Systems* **213**:106706 DOI 10.1016/j.knosys.2020.106706.

Kai-yuan L. 2022. Patterns recognition of unsafe behavior in chemical laboratory based on C3D. *Information Technology and Network Security* **41(3)**:71–77.

Kim S, Ahn D, Ko BC. 2023. Cross-modal learning with 3D deformable attention for action recognition. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Piscataway: IEEE, 10265–10275.

Liu H, Cai K, Li P, Qian C, Zhao P, Wu X. 2023. REDRL: a review-enhanced deep reinforcement learning model for interactive recommendation. *Expert Systems with Applications* **213(3)**:118926 DOI 10.1016/j.eswa.2022.118926.

Liu R, Shen J, Wang H, Chen C, Cheung SC, Asari VK. 2021. Enhanced 3D human pose estimation from videos by using attention-based neural network with dilated convolutions. *International Journal of Computer Vision* **129(5)**:1596–1615 DOI 10.1007/s11263-021-01436-0.

Liu F, Tang R, Li X, Zhang W, Ye Y, Chen H, He X. 2020. State representation modeling for deep reinforcement learning based recommendation. *Knowledge-Based Systems* **205**:106170 DOI 10.1016/j.knosys.2020.106170.

Peng Y, Siet S, Ilkhomjon S, Kim DY, Park DS. 2024. Integration of deep reinforcement learning with collaborative filtering for movie recommendation systems. *Applied Sciences* **14(3)**:1155 DOI 10.3390/app14031155.

Prakash C, Kumar R, Mittal N. 2018. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artificial Intelligence Review* **49**:1–40 DOI 10.1007/s10462-016-9514-6.

Ren B, Liu M, Ding R, Liu H. 2024. A survey on 3D skeleton-based action recognition using learning method. *Cyborg and Bionic Systems* **5**:100 DOI 10.34133/cbsystems.0100.

Sampaio T, Oliveira JP, Marinho DA, Neiva HP, Morais JE. 2024. Applications of machine learning to optimize tennis performance: a systematic review. *Applied Sciences* **14(13)**:5517 DOI 10.3390/app14135517.

Skublewska-Paszkowska M, Powroznik P, Lukasik E, Smolka J. 2024. Tennis patterns recognition based on a novel tennis dataset-3DTennisDS. *Advances in Science and Technology. Research Journal* **18(6)**:159–176 DOI 10.12913/22998624/191264.

Sohafi-Bonab J, Aghdam MH, Majidzadeh K. 2023. DCARS: deep context-aware recommendation system based on session latent context. *Applied Soft Computing* **143(6)**:110416 DOI 10.1016/j.asoc.2023.110416.

Tu Y, Lin S, Qiao J, Zhuang Y, Zhang P. 2022. Alzheimer's disease diagnosis via multimodal feature fusion. *Computers in Biology and Medicine* **148**:105901 DOI 10.1016/j.compbiomed.2022.105901.

Wang T, Hou B, Li J, Shi P, Zhang B, Xu M, Liu K, Snoussi H. 2023. TASTA: text-assisted spatial and temporal attention network for video question answering. *Advanced Intelligent Systems* **5(4)**:2200131 DOI 10.1002/aisy.202200131.

Wang S, Wu X, Lai W, Yao J, Gou X, Ye H, Yi J, Cao D. 2025. Rehabilitation evaluation method and application for upper limb post-stroke based on improved DTW. *Biomedical Signal Processing and Control* **106**:107775 DOI 10.1016/j.bspc.2025.107775.

Wang JH, Wu YT, Wang L. 2021. Predicting implicit user preferences with multimodal feature fusion for similar user recommendation in social media. *Applied Sciences* **11(3)**:1064 DOI 10.3390/app11031064.

Wei H, Zheng G, Gayah V, Li Z. 2021. Recent advances in reinforcement learning for traffic signal control: a survey of models and evaluation. *ACM SIGKDD Explorations Newsletter* **22(2)**:12–18 DOI 10.1145/3447556.3447565.

Xin X, Karatzoglou A, Arapakis I, Jose JM. 2022. Supervised advantage actor-critic for recommender systems. In: *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*. New York: ACM, 1186–1196.

Yang R, Li H, Huang H. 2024. Multisource information fusion considering the weight of focal element's beliefs: a Gaussian kernel similarity approach. *Measurement Science and Technology* **35(2)**:025136 DOI 10.32388/3c9rw8.

Zhang R, Wang Y, Li Z, Ding F, Wei C, Liu Y, Wang J, Wu M. 2025. Online adaptive keypoint extraction for visual odometry across different scenes. *IEEE Robotics and Automation Letters* **10(7)**:7539–7546 DOI 10.1109/LRA.2025.3575644.

Zhao X, Wang T, Li Y, Zhang B, Liu K, Liu D, Zhou F, Hu B, Snoussi H. 2024. Target-driven visual navigation by using causal intervention. *IEEE Transactions on Intelligent Vehicles* **9(1)**:1294–1304 DOI 10.1109/TIV.2023.3288810.

Zhu S, Yu K, Chi Y, Gong Y. 2007. Combining content and link for classification using matrix factorization. In: *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York: ACM, 487–494.