

Deep Learning-Based Badminton Swing Action Recognition via Quaternion Gated Recurrent Unit and Dynamic Attention

Longyi Yuan¹, Wei Wang^{1*}, Yi Chen¹

¹ School of Life Health Information Science and Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China

*Correspondence: Wei Wang

Received: 2 April 2026; Accepted: 28 April 2026; Published: 20 May 2026

Abstract:

Inertial sensor-based recognition of badminton swings faces significant challenges, including cumulative sensor drift, attitude representation singularities, and the difficulty of balancing global motion trends with local explosive features. To address these issues, this paper constructs a high-precision kinematic trajectory reconstruction framework and a lightweight recognition architecture based on physical priors. Following rigorous sensor calibration, 22 volunteers (20 advanced amateurs and 2 beginners) were recruited to collect 5,940 valid movement samples covering nine core technical actions. The proposed model utilizes 4-dimensional quaternion components as backbone input features and 1-dimensional angular velocity magnitude as a branch modulation parameter. A Quaternion Gated Recurrent Unit (QvGRU) is employed to extract features directly in the quaternion domain, reducing the number of parameters while maintaining the integrity of the 3D attitude space. The extracted features are then processed by an Angular Velocity Modulated (AVM) dual dynamic high-low frequency attention module to adapt to drastic signal bandwidth fluctuations. Experimental results demonstrate that the proposed model achieves a recognition accuracy of 98.70% on the correct action test set and maintains 95.74% on the beginner test set, proving robust generalization capability. Furthermore, the model is highly lightweight, with only 28.4K parameters and a computational load of 4.55 MFLOPs.

Keywords:

Inertial Sensor-based Recognition, Badminton Swings, Quaternion Gated Recurrent Unit (QvGRU), Angular Velocity Modulated (AVM)

1. Introduction

The precise recognition of badminton swing actions is a critical prerequisite for optimizing technical performance and preventing sports injuries [1,2]. Traditional action analysis relies heavily on subjective coaching experience. Recently, wearable inertial measurement units (IMUs) have demonstrated significant advantages over optical motion capture (OMC) and surface electromyography (sEMG) due to their low cost, high portability, and environmental independence [3,4].

However, current IMU-based recognition methods face several limitations [5,6]. Complex neural networks are often too computationally heavy for edge deployment, and standard real-valued neural networks tend to fragment the highly coupled 3D spatial properties of inertial data [7,8].

To resolve these limitations, we propose the QvGRU-AVMHiLo model. By utilizing quaternions and angular velocity magnitude as explicit physical priors, and by employing a highly efficient quaternion-based recurrent architecture, the model strikes an optimal balance between high precision and low computational cost.

2. Methodology

2.1. Trajectory Reconstruction and Dataset

To establish a physically consistent input, a rigorous trajectory reconstruction framework was adopted. Motion data was acquired using a wearable 9-axis Inertial Measurement Unit (IMU) consisting of a 3-axis gyroscope, a 3-axis accelerometer, and a 3-axis magnetometer. To accurately capture the highly dynamic and explosive nature of badminton swings, the sensor parameters were configured with a gyroscope range of $\pm 2000^\circ/\text{s}$, an accelerometer range of $\pm 16\text{g}$, and a magnetometer range of ± 8 Gauss. Continuous motion signals were recorded at a stable sampling rate of 60 Hz. Subsequently, acceleration and gyroscope calibrations were performed to suppress hardware errors.

In this study, a total of 22 volunteers were recruited, including 18 males and 4 females. (Age of 22 ± 4 years, height of 170.5 ± 15.5 cm, and weight of 62 ± 14 kg. Comprising 20 advanced amateurs and 2 beginners) to execute nine core badminton actions, including forehand clear, smash, and backhand shots. Using a sliding window triggered by the peak angular velocity magnitude, we successfully isolated 5,940 standardized swing samples (5,400 from advanced amateurs and 540 from beginners).

After standardization, the 5,400 swing action samples from advanced amateurs were partitioned into a training set, a validation set, and Test Set A at a ratio of 8:1:1. The model was trained exclusively using the motion data from advanced amateurs, while the 540 swing action samples from beginners served as Test Set B to evaluate the model's generalization capability.

2.2. QvGRU-AVMHiLo Architecture

The proposed model abandons the direct stacking of raw signals. Instead, it constructs a 5-dimensional feature vector: a 4D quaternion for spatial rotation and a 1D angular velocity magnitude for dynamic force. The proposed model architecture primarily consists of three core modules: the feature extraction section, the average pooling module, and the final classification layer. Guided by physical priors, the core design of the model integrates temporal modeling in the quaternion domain with dynamic frequency feature extraction, thereby achieving precise classification of the nine types of badminton swing actions.

The feature extraction section serves as the backbone of the entire network and is composed of two key components: first, a Quaternion Gated Recurrent Unit (QvGRU) module executes operations natively within the quaternion space to preserve the structural integrity of the sensor rotation data while preliminarily capturing temporal dependencies; subsequently, an Angular Velocity Modulated dual dynamic high-low

frequency attention (AVMHiLo) module refines these features, leveraging the angular velocity magnitude as a physical prior to dynamically allocate computational resources for simultaneously capturing high-frequency local explosive features and low-frequency global trend features.

Following feature extraction, the model enters the average pooling stage, where dimensional permutation and adaptive average pooling are sequentially performed to compress the multi-dimensional temporal features into a global feature vector. Finally, this vector is forwarded to a classifier containing multiple linear layers, Batch Normalization, and a Dropout mechanism, where a Softmax function outputs the probability distribution across all swing categories, designating the class with the highest probability as the final recognition result. In the following sections, the design details and implementation logic of the two core components, QvGRU and AVMHiLo, will be explored in depth. (Figure 1)

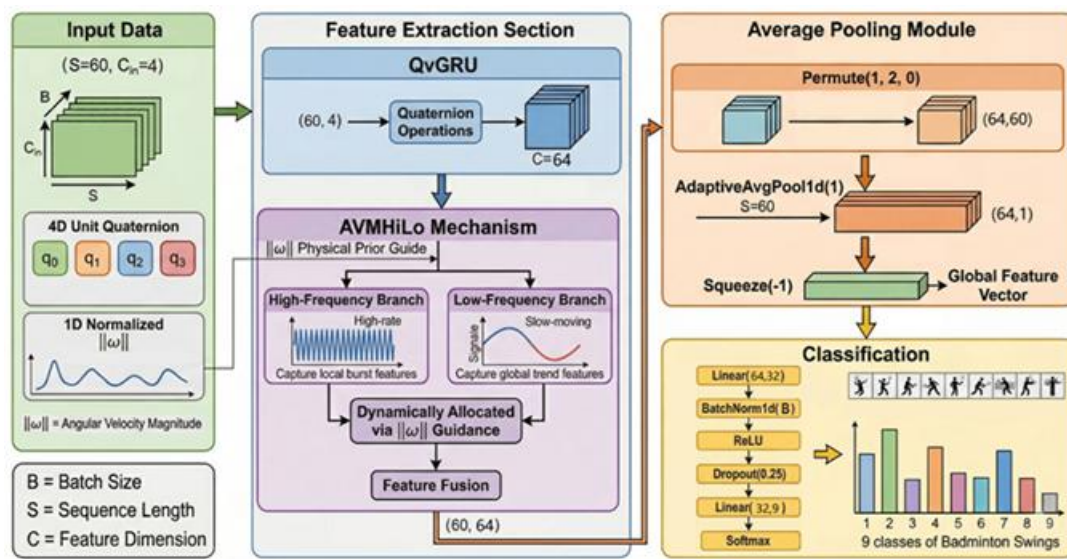


Figure 1. Architecture of the QvGRU-AVMHiLo model.

2.2.1. Quaternion Gated Recurrent Unit (QvGRU)

QvGRU processes the 4D input natively using Hamilton products. This perfectly preserves the internal physical coupling of 3D rotations. Crucially, the weight-sharing nature of the quaternion neural network reduces the parameter count to one-quarter of that of a real-valued GRU[9].

2.2.2. Angular Velocity Modulated Attention (AVMHiLo)

To handle the bandwidth mutation between the preparation phase (low frequency) and the hitting phase (high frequency), we designed a dual dynamic attention module. A Gumbel dynamic routing mechanism adaptively allocates computational resources to high- and low-frequency attention branches based on the instantaneous angular velocity. (Figure 2)

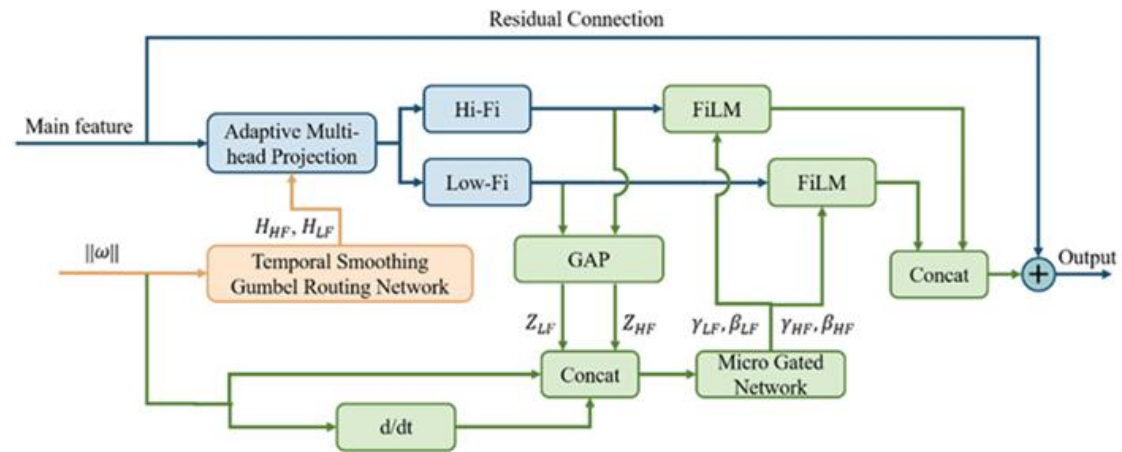


Figure 2. AVMHiLo architecture diagram.

3. Experiments and Results

3.1. Experimental Setup

The hardware environment for the experiments in this chapter utilized an Intel Core i5-10500 CPU and an NVIDIA GeForce GTX 1050 Ti GPU. The experiments were conducted using PyTorch 2.0.0 as the deep learning framework, with Python 3.9 (PyCharm 2025.1.3.1) serving as the development language and environment. GPU acceleration was supported by CUDA 11.8.

During the training process, random noise addition and angular rotation were applied as data augmentation techniques. This preprocessing approach effectively mitigates the out-of-distribution (OOD) issues inherent in IMU data, thereby endowing the model with enhanced generalization capabilities.

The training process consisted of 100 epochs, with a batch size of 32. Cross-Entropy loss was utilized as the loss function. Furthermore, the Adam optimizer was employed; by integrating the adaptive learning rate of AdaGrad with the momentum concept of RMSProp, it enables more efficient model parameter updates while mitigating overfitting.

A validation phase was conducted at the conclusion of each training epoch. In this chapter, a dynamic learning rate scheduling strategy based on validation metrics was adopted. Unlike fixed-epoch decay methods (e.g., StepLR), this strategy dynamically adjusts the learning rate according to fluctuations in validation metrics (loss or accuracy). When the model ceases to improve on the validation set, the learning rate is reduced to facilitate escaping local optima. The initial learning rate was set to 10^{-3} as the foundational optimization step size. Specifically, if the calculated validation loss failed to decrease for three consecutive epochs, a decay mechanism was triggered, dividing the learning rate by 2 and resetting the patience counter to 0. This strategy guarantees rapid convergence in the early stages of training while achieving superior precision and generalization through fine-grained adjustments in the later stages. It represents a highly stable and flexible learning rate adjustment method, making it exceptionally well-suited for the training requirements of complex models.

3.2. Recognition Accuracy

The proposed QvGRU-AVMHiLo model achieved an outstanding recognition accuracy of 98.70% on the advanced amateurs test set. Furthermore, when evaluated

on non-standard actions performed by beginners, the model maintained an accuracy of 95.74%, demonstrating exceptional feature robustness and cross-subject generalization capability. (Table 1)

Table 1. Recognition results of the QvGRU-AVMHiLo model.

Swing Action	Advanced Amateur Test Set (Test Set A)			Beginner Test Set (Test Set B)		
	Precision / %	Recall / %	F1 Score / %	Precision / %	Recall / %	F1 Score / %
Forehand Clear	100.00	100.00	100.00	94.74	90.00	92.31
Forehand Drive	100.00	98.33	99.16	91.94	95.00	93.44
Forehand Lift	98.33	98.33	98.33	98.36	100.00	99.17
Forehand Smash	98.28	95.00	96.61	91.53	90.00	90.76
Forehand Slice Drop	95.24	100.00	97.56	93.33	93.33	93.33
Forehand Slide Drop	100.00	100.00	100.00	98.36	100.00	99.17
Backhand Clear	100.00	96.67	98.31	98.31	96.67	97.48
Backhand Drive	98.36	100.00	99.17	96.77	100.00	98.36
Backhand Lift	100.00	100.00	100.00	98.31	96.67	97.48

3.3. Model Efficiency

Efficiency evaluations confirm the lightweight nature of our approach compared to standard baseline models. The results are summarized in the table below:

Table 2. Comparative results of recognition accuracy.

Model	Accuracy A (%)	Accuracy B (%)	Parameters (K)	MFLOPs
SA-ConvLSTM [10]	96.85	91.20	55.5	7.15
Ours	98.70	95.74	28.4	4.55

As illustrated in the table, the proposed model demonstrates superior performance across all evaluated metrics compared to the baseline SA-ConvLSTM model. Specifically, our approach achieves an Accuracy A of 98.70% and an Accuracy B of 95.74%, outperforming the baseline by margins of 1.85% and 4.54%, respectively. The more pronounced improvement in Accuracy B indicates that our model possesses stronger generalization capabilities and robustness. Furthermore, these performance gains are achieved with a significantly reduced computational footprint. Requiring only 28.4K parameters and 4.55 MFLOPs, our model reduces the parameter count by nearly 49% and computational complexity by over 36% compared to SA-ConvLSTM. These results confirm that the proposed architecture successfully strikes an optimal balance between high recognition accuracy and exceptional computational efficiency.

4. Conclusions

This paper presents a low-cost, highly robust badminton swing recognition solution. By addressing cumulative drift through rigorous kinematic reconstruction and proposing the QvGRU-AVMHiLo network, we effectively capture the dynamic and

spatial complexities of badminton swings. The physical prior-guided architecture ensures high accuracy while maintaining a remarkably low computational footprint of 4.55 MFLOPs. However, this study has certain limitations. Specifically, the model exhibits a slight performance degradation when processing highly non-standard actions, such as the beginner's forehand smash which yielded an F1 score of 90.76%, indicating that extreme kinematic variability among novices remains a challenge. Furthermore, while the theoretical parameter count is minimized, real-world deployment constraints such as continuous battery consumption and strict real-time inference latency on edge devices require further empirical validation. Future work will focus on expanding the dataset diversity and exploring multimodal fusion strategies, such as integrating IMU kinematics with surface electromyography (sEMG) for continuous motion intention decoding. Additionally, we aim to extend this highly efficient edge-computing architecture beyond ubiquitous sports analytics, exploring its potential as a wearable algorithmic solution for medical devices in motor rehabilitation tracking and clinical functional assessment.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- [1] Tajik, R.; Dhahbi, W.; Fadaei, H, et al. Muscle synergy analysis during badminton forehand overhead smash: integrating electromyography and musculoskeletal modeling. *Frontiers in Sports and Active Living*, 2025, 7, 1596670.
- [2] Ni, J.; Tang, H.; Haque, S.T, et al. A survey on multimodal wearable sensor-based human action recognition. arXiv preprint arXiv:2404.15349, 2024.
- [3] Sanchez, A, et al. TinyML for Real-time Motion Analysis: A Survey of On-device Inference. *IEEE Internet of Things Journal*. 2025, 12(3), 4501-4518.
- [4] Sola, J. Quaternion kinematics for the error-state Kalman filter. arXiv preprint arXiv:1711.02508, 2017.
- [5] Gayathri Devi, S.; Siva Harish, R.V.; Nalini, N, et al. Advancing human activity recognition with quaternion-based recurrent neural networks. *Automatika*, 2025, 66(3), 411-430.
- [6] Xia, K.; Wang, H.; Xu, M, et al. Racquet sports recognition using a hybrid clustering model learned from integrated wearable sensor. *Sensors*, 2020, 20(6), 1638.
- [7] Wu, Y.; Tao, K.; Chen, Q, et al. A comprehensive analysis of the validity and reliability of the perception neuron studio for upper-body motion capture. *Sensors*, 2022, 22(18), 6954.

- [8] Van der Kruk, E.; Reijne, M.M. Accuracy of human motion capture systems for sport applications; state-of-the-art review. *European journal of sport science*, 2018, 18(6), 806-819.
- [9] Liu, J.; Liang, B. An action recognition technology for badminton players using deep learning. *Mobile Information Systems*, 2022, 2022(1), 3413584
- [10] Deng, J.; Zhang, S.; Ma, J. Self-Attention-Based deep Convolution LSTM framework for Sensor-Based badminton activity recognition. *Sensors*, 2023, 23(20), 8373.



© 2026 by the author(s); licensee International Technology and Science Publications (ITS), this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (<http://creativecommons.org/licenses/by/4.0/>)