Adaptive IMU Dynamic Calibration For Intelligent Perception Of High-Dynamic Sports-Like Motions

Dong Xiaoou[1], Guan Xinru[2], Feng Yongxu[1], and Lu Jiawei[1*]

[1] School of Physical Education, Jiamusi University, Jiamusi 154007, China
[2] College of Information and Electronic Technology, Jiamusi University, Jiamusi 154007, China
[*] Corresponding author. E-mail:18745521387@163.com

Received: Apr. 26, 2026; Accepted: May. 18, 2026

In intelligent perception of high-dynamic sports-like motions, inertial measurement units (IMUs) are susceptible to sensor slippage, mounting misalignment, and global drift during prolonged high-intensity motion, which progressively degrades the validity of one-time static calibration. To address this issue, this paper proposes an adaptive IMU dynamic calibration method for high-dynamic sports-like motion perception. The proposed framework takes rotation matrices and acceleration sequences from six body-worn IMUs as input and constructs an online calibration architecture based on a Transformer Encoder backbone. Through sliding-buffer inference, motion-diversity-triggered updating, and dual-branch parameter estimation, the method recursively updates the drift matrix R_{DG} and the mounting offset matrix R_{BS}. Compared with static-only calibration, the proposed method is better suited for continuous correction under motion-intensive patterns such as turning, punching-like arm swings, jumping-like bursts, and abrupt direction changes. Experimental results indicate stable convergence and effective parameter estimation. The experimental protocol includes static baseline comparison, original dynamic calibration reproduction, validation on a high-dynamic sports-like proxy subset constructed from the TIC test set, trigger ablation, and downstream proxy action recognition verification. These results demonstrate that the proposed framework provides a practical solution for online calibration in intelligent sensing systems for motion-intensive sports-like movements.

Keywords: high-dynamic sports-like motions; IMU dynamic calibration; Transformer; online calibration; intelligent perception

http://dx.doi.org/10.6180/jase.202609_32.056

1. Introduction

Sports movement monitoring and intelligent perception aim to quantify body motion patterns, action structures, and training performance, so as to provide scientific support for training load management, technique optimization, and injury prevention[1]. With the rapid development of wearable sensing technologies, inertial measurement unit (IMU)-based motion analysis has become one of the mainstream solutions in sports monitoring due to its low cost, portability, and real-time capability[2, 3].

However, in high-dynamic sports-like motions, the limitations of IMU-based sensing become much more pronounced. First, IMUs mainly capture local motion information around the sensor-wearing location, making it difficult to fully characterize coordinated full-body actions[4]. Second, in long-term movement sequences, static calibration parameters gradually become invalid due to strap looseness, body-surface deformation, and relative displacement between the sensor and the body segment[5]. Third, high-intensity motion patterns such as turning, abrupt stopping, punching-like arm swings, and jumping-like bursts accelerate the accumulation of global drift, thereby reducing

the reliability of downstream pose estimation and motion analysis. Therefore, traditional static-only calibration strategies are often insufficient for intelligent perception in high-dynamic sports-like motion scenarios[6].

To address this problem, this paper proposes an adaptive IMU dynamic calibration method for intelligent perception of high-dynamic sports-like motions. Instead of relying on one-time static calibration before the experiment, the proposed method continuously estimates calibration parameters during motion by leveraging multi-IMU temporal observations. In addition, a motion-diversity-triggered mechanism is introduced to determine when calibration updating should be activated, thus preventing unreliable corrections in low-information windows. The overall research route is divided into two stages: first, the original Transformer Encoder-based IMU Calibrator (TIC) baseline is reproduced to verify the complete dynamic calibration pipeline; second, the framework is extended to high-dynamic sports-like proxy scenarios to evaluate its effectiveness in sport-relevant motion segments.

The main contributions of this work are summarized as follows. (1) A Transformer Encoder-based dual-branch online dynamic calibration framework is constructed to jointly estimate IMU global drift and mounting offset. (2) A motion-diversity-triggered updating strategy is introduced to adapt calibration updates to the temporal characteristics of high-dynamic sports-like motions. (3) An experimental protocol is designed around static baseline comparison, validation on a sports-like proxy subset, trigger ablation, and downstream proxy action recognition verification, thereby establishing a complete research route for high-dynamic sports-like motion sensing.

2. Related work

2.1. Wearable Sports Monitoring and Intelligent Perception

Wearable sensing systems have been widely used in running monitoring, strength-training analysis, exercise recognition, and sports performance evaluation[7–9]. Related research has reviewed the evolution from manual observation and video-based analysis to intelligent wearable sensing, and has discussed activity recognition and performance assessment as two major research directions[10, 11]. Nevertheless, real-world sports scenarios are characterized by large motion diversity, unstable sensor wearing conditions, and substantial inter-subject variability. As a result, many existing systems still suffer from limited cross-subject generalization and insufficient robustness in complex movement patterns.

2.2. IMU-Based Sports Sensing and Its Limitations

IMUs have been widely adopted in sports monitoring because of their compactness, low power consumption, and high sampling capability[12, 13]. However, their sensing ability is inherently local and highly dependent on sensor placement and orientation consistency[14]. Existing studies have also shown that IMUs are limited in cross-body information perception, integration drift, and low-dynamic movement discrimination[15, 16]. In the present work, these limitations are further considered from the perspective of dynamic calibration. Specifically, two major error sources are addressed: global reference drift and body-segment-related mounting offset[17, 18]. If such errors are not corrected online, they directly degrade the accuracy of pose estimation and high-dynamic sports-like motion analysis[19].

2.3. Dynamic Calibration and Temporal Modeling

Recent calibration research is gradually shifting from manually designed geometric correction strategies to deep-learning-based temporal modeling methods. Compared with conventional convolutional networks, Transformer Encoder architectures are better suited to capturing long-range temporal dependencies in multivariate motion sequences[20, 21]. In the present study, a Transformer Encoder-based framework is incorporated into IMU dynamic calibration together with a motion-diversity-triggered update mechanism to improve calibration stability under high-dynamic sports-like motions[22, 23].

3. Method

3.1. Overall Model Architecture

The proposed adaptive IMU dynamic calibration framework for high-dynamic sports-like motion perception is illustrated in Fig. 1. Unlike conventional one-time static calibration strategies, the proposed framework performs online recursive calibration during motion and consists of five tightly coupled components: a wearable IMU sensing module, a sensor data acquisition and preprocessing module, a motion-diversity-triggered update module, a Transformer Encoder-based dual-branch calibration network, and a calibrated output module for downstream intelligent perception.

Specifically, six IMUs are mounted on the head, left forearm, right forearm, left shank, right shank, and pelvis, respectively, so as to provide sparse yet informative body motion observations. The raw rotational matrices and acceleration signals are continuously collected and written into a sliding buffer. Once the buffered sequence reaches

Table 1. Notation and key parameter definitions used in this study.

No.	Symbol	Description	Value	Type
1	C_{IMU}	Number of IMUs	6	Sensor
2	$\tilde{R}_t^{(i)}$	Raw rotational observation of the i-th IMU	3×3 matrix	Input
3	$\tilde{a}_t^{(i)}$	Raw acceleration observation of the i-th IMU	3-D vector	Input
4	X	Model input sequence tensor	256×72	Input
5	$R_{DG,t}^{(i)}$	Drift compensation matrix	Dynamic	Calibration
6	$R_{BS,t}^{(i)}$	Mounting offset compensation matrix	Dynamic	Calibration
7	ΔR_{DG}	Increment of drift correction	–	Calibration
8	ΔR_{BS}	Increment of offset correction	–	Calibration
9	D	Motion diversity vector	6-D vector	Trigger
10	τ_{DG}/τ_{BS}	Per-IMU drift / offset update thresholds	6-D vectors	Trigger
11	L_{buf}	Online buffer length	512 frames	Setting
12	L_{seq}	Training sequence length	256 frames	Setting
13	f_s/f_{in}	Raw / input sampling frequency	60 / 30 Hz	Setting
14	d_{model}	Transformer hidden dimension	256	Model
15	n_{head}	Number of attention heads	8	Model
16	d_{ff}	Feedforward hidden dimension	512	Model
17	J / L_{train}	Conceptual objective / implemented training loss	Weighted sum / branch-wise MSE	Objective / Loss

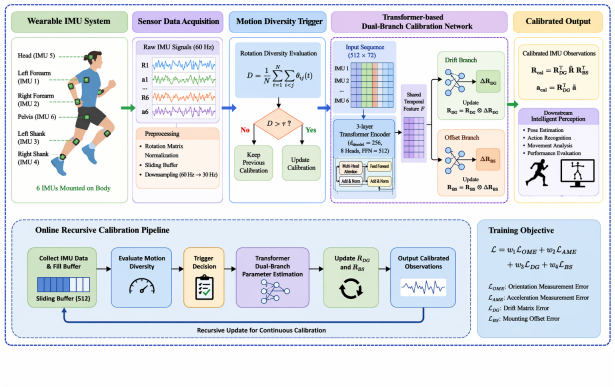

Fig. 1. System Architecture

the predefined window length, the system evaluates the motion diversity of the current segment. If the motion information is sufficiently rich, the buffered sequence is fed into the Transformer Encoder-based dual-branch calibration network to estimate the drift correction and mounting offset correction. The updated calibration parameters are then recursively applied to produce calibrated observations, which are further used for downstream tasks such as pose estimation, action recognition, movement analysis, and performance evaluation. Compared with multimodal multi-task perception frameworks, the present framework focuses on an online dynamic calibration pipeline centered on motion-triggered parameter updating and recursive correction.

3.1.1. Notation and Parameter Definitions

Key symbols and parameters are defined consistently throughout the paper, and their values in the methodology section correspond to the experimental settings, as shown in Table 1.

3.2. Wearable IMU System and Sensor Data Acquisition

As shown in Fig. 1, the framework starts from a wearable IMU sensing system. Six IMUs are mounted on representative upper- and lower-body locations, including the bilateral forearms, bilateral shanks, the head, and the pelvis. This sparse body-mounted configuration is intended to balance deployment convenience and motion observability in high-dynamic sports-like motion scenarios. The choice of these six positions allows the framework to simultaneously capture upper-limb, lower-limb, and trunk-related motion variations.

The raw IMU signals include rotational observations and acceleration observations. During online operation, all sensor streams are continuously recorded and stored in a sliding buffer. Before entering the calibration network, the signals are preprocessed through rotational matrix organization, normalization, buffering, and temporal downsampling. This preprocessing module converts raw sensor streams into temporally aligned and normalized sequences suitable for robust online calibration.

The online calibration process uses a sliding buffer to maintain temporal consistency between raw IMU observations and the model input. At each time step, the raw inertial measurement unit (IMU) data, including acceleration vectors and rotation matrices from six sensors, are first corrected using the current calibration state and then appended to the buffer in chronological order. The buffer size is fixed to 512 frames. When the number of stored frames is smaller than 512, no dynamic calibration is performed.

Once the buffer contains at least 512 frames, the most recent 512-frame segment is extracted for model inference.

Since the raw sequence is recorded at 60 Hz while the Transformer Encoder is trained with 30 Hz input, the extracted segment is uniformly downsampled by a factor of two. Specifically, every second frame is selected from the 512-frame segment, resulting in a 256-frame input sequence:

$$X^{30\text{Hz}} = \{X_{t-511}, X_{t-509}, \ldots, X_{t-1}\} \tag{1}$$

where t denotes the current frame index. This operation preserves the chronological order of the original sequence and ensures that the model input is temporally aligned with the latest raw observations.

The dynamic calibration module is triggered at a fixed interval, e.g., every 1 s. At each trigger time, the model estimates calibration increments from the downsampled 256-frame sequence. If the motion diversity condition is satisfied, the calibration state is updated and applied to subsequent IMU observations. Therefore, the raw 60 Hz stream, the 512-frame sliding buffer, and the 30 Hz model input are aligned through deterministic temporal indexing, ensuring reproducible online calibration.

3.3. Motion Diversity Trigger

A key innovation of the proposed framework lies in the motion-diversity-triggered update mechanism. In high-dynamic sports-like motions, not all temporal windows are equally informative for calibration. If calibration parameters are updated indiscriminately across low-motion or low-information segments, the system may introduce unstable corrections. To address this issue, the current framework introduces a motion diversity evaluation block that quantifies the richness of rotational variation within the buffered sequence.

The motion diversity score D is designed to quantify the richness of IMU orientation changes within the current online calibration window. It is used as a gating criterion to determine whether the dynamic calibration parameters should be updated.

For the i-th IMU, the rotation matrix sequence within the current window is denoted as

$$\{R_{i,t}\}_{t=1}^{N}, \tag{2}$$

where $N = 256$. This corresponds to a 512-frame online buffer downsampled from 60 Hz to 30 Hz. Each rotation matrix $R_{i,t}$ is first converted into XYZ Euler angles in degrees:

$$e_{i,t} = (\alpha_{i,t}, \beta_{i,t}, \gamma_{i,t}). \tag{3}$$

The Euler angles are then discretized using a bin width of 15°:

$$b_{i,t}^{x} = \text{clip}\left(\left\lfloor \frac{\alpha_{i,t}}{15} \right\rfloor + 12, 0, 23\right), \tag{4}$$

$$b_{i,t}^{y} = \text{clip}\left(\left\lfloor \frac{\beta_{i,t}}{15} \right\rfloor + 6, 0, 11\right), \tag{5}$$

$$b_{i,t}^{z} = \text{clip}\left(\left\lfloor \frac{\gamma_{i,t}}{15} \right\rfloor + 12, 0, 23\right), \tag{6}$$

where $\text{clip}(\cdot)$ constrains the index to the specified range. The three discretized angular components are further mapped to a one-dimensional orientation-state index:

$$q_{i,t} = b_{i,t}^{x} + 24b_{i,t}^{y} + 24 \times 12 b_{i,t}^{z}. \tag{7}$$

Thus, the orientation space of each IMU is partitioned into $24 \times 12 \times 24$ discrete cells. The motion diversity score of the i-th IMU is defined as the number of distinct orientation cells visited within the current window:

$$D_i = \sum_{q=0}^{24\times12\times24-1} \mathbb{I}\left(\sum_{t=1}^{N} \mathbb{I}(q_{i,t} = q) > 0\right). \tag{8}$$

The final motion diversity vector is

$$D = (D_1, D_2, D_3, D_4, D_5, D_6), \tag{9}$$

corresponding to the six IMUs mounted on the left forearm, right forearm, left shank, right shank, head, and pelvis, respectively.

The parameter settings are selected according to the online inference configuration and the stability requirement of dynamic calibration. The 512-frame buffer covers approximately 8.53 s of motion at the original 60 Hz sampling rate, while downsampling to 30 Hz ensures consistency with the input setting of the TIC network. The 15° bin width balances sensitivity and robustness: a smaller bin width would make D more sensitive to sensor noise and small orientation fluctuations, whereas a larger bin width would reduce its ability to distinguish informative high-dynamic motion patterns.

The trigger thresholds τ_{DG} and τ_{BS} are defined as IMU-specific motion diversity thresholds for accepting dynamic calibration updates. Here, τ_{DG} is associated with the drift correction matrix R_{DG}, and τ_{BS} is associated with the mounting-offset correction matrix R_{BS}. In the implementation, they are set as:

$$\tau_{DG} = [10, 10, 10, 10, 10, 0], \tag{10}$$

$$\tau_{BS} = [30, 50, 30, 30, 25, 15]. \tag{11}$$

where the six elements correspond to the left forearm, right forearm, left shank, right shank, head, and pelvis IMUs, respectively. During online calibration, the drift correction matrix R_{DG} of the i-th IMU is updated only when $D_i > \tau_{DG,i}$, while the mounting-offset correction matrix R_{BS} is updated only when $D_i > \tau_{BS,i}$. If the condition is not satisfied, the corresponding correction increment is replaced by an identity matrix to avoid unreliable updates from low-information motion segments.

The threshold values were selected according to the motion diversity distribution observed in the online calibration windows and the stability requirement of recursive calibration. Specifically, τ_{DG} is set relatively low because global drift correction can benefit from moderate orientation variation, whereas τ_{BS} is set higher because estimating sensor-to-body mounting offset requires richer relative orientation coverage. The pelvis IMU uses a lower or zero drift threshold because it provides the global reference for ego-motion regularization and usually exhibits smaller orientation diversity than limb-mounted IMUs.

The IMU-specific design reflects the different motion characteristics of body segments. Forearm IMUs generally show larger and more complex orientation changes during sports actions, so their mounting-offset thresholds are higher. Shank and head IMUs use moderate thresholds because their motion diversity is lower or more constrained. The pelvis IMU uses a lower threshold due to its relatively stable motion pattern and its role as the reference segment.

The validity of these thresholds was verified through the ablation comparison between the original TIC configuration and the w/o Trigger setting. When the motion-diversity gating was removed, all predicted calibration increments were accepted regardless of D, which increased the risk of updates from static or low-excitation windows. The original threshold-based strategy produced more stable calibration behavior, indicating that the selected τ_{DG} and τ_{BS} effectively suppress unreliable updates while preserving informative high-dynamic calibration opportunities.

3.4. Transformer Encoder-based Dual-Branch Calibration Network

Once the trigger condition is satisfied, the buffered sequence is fed into the Transformer Encoder-based dual-branch calibration network. This module constitutes the core of the proposed method. The online buffer contains 512 raw frames, which are downsampled to a 256-frame model input. Therefore, the input tensor fed into the Transformer Encoder has a size of 256×72, where 256 denotes the model input sequence length and 72 denotes the concatenated feature dimension of all six IMUs. More specifically, each IMU provides a 3D acceleration vector and a 3×3 rotation matrix in each frame. Therefore, the frame-wise input dimension is given by

$$6 \times (3 + 9) = 72. \tag{12}$$

The TIC network adopts a 3-layer Transformer Encoder as the temporal feature extraction backbone. Each encoder layer consists of a multi-head self-attention module and a position-wise feedforward network. In this study, the model dimension is set to $d_{model} = 256$, the number of attention heads is set to 8, and the feedforward dimension is set to $d_{ff} = 512$.

These settings are chosen to balance temporal modeling capacity and computational efficiency. The 3-layer encoder provides sufficient depth to capture short- and medium-term temporal dependencies in the 256-frame IMU sequence, while avoiding excessive model complexity for online calibration. The 8-head attention mechanism allows the model to attend to different motion patterns across IMU channels and time steps, such as limb swing, global rotation, and local sensor perturbation. The model dimension $d_{model} = 256$ ensures that the 72-dimensional IMU input can be projected into a richer latent representation, while $d_{ff} = 512$ expands the feature space within each encoder block to improve nonlinear representation learning.

This configuration also keeps the network lightweight enough for online inference. Therefore, the selected number of layers, attention heads, and feedforward dimension represent a compromise between calibration accuracy, temporal representation capability, and computational cost.

Based on the shared temporal feature representation F, the network is divided into two estimation branches: a drift branch and a mounting-offset branch. The drift branch predicts the global drift increment, whereas the mounting-offset branch predicts the local sensor-to-body offset increment.

Each branch outputs a 36-dimensional vector composed of the 6D rotation representations of six IMUs, i.e.,

$$36 = 6 \times 6. \tag{13}$$

Accordingly, the drift branch outputs

$$\hat{d}^{DG} \in \mathbb{R}^{36}, \tag{14}$$

and the mounting-offset branch outputs

$$\hat{d}^{BS} \in \mathbb{R}^{36}. \tag{15}$$

For the j-th IMU, the network predicts two 6D rotation vectors,

$$\hat{d}_j^{DG} \in \mathbb{R}^6, \qquad \hat{d}_j^{BS} \in \mathbb{R}^6. \tag{16}$$

These 6D rotation representations are then converted into 3×3 rotation increment matrices, denoted as ΔR_j^{DG} and ΔR_j^{BS}, which are used for recursive state updating during online inference. Therefore, the total output dimensionality of the dual-branch head is 72, consisting of 36 dimensions for drift estimation and 36 dimensions for mounting-offset estimation. The 6D rotation representation is adopted to

alleviate the discontinuity and instability issues associated with Euler-angle- or quaternion-based parameterizations, thereby providing a more stable representation for learning rotational increments. The use of two separate branches is motivated by the physical distinction between global reference drift and local sensor-to-segment mounting deviation. By decoupling these two sources of error in a dual-branch manner, the framework improves the interpretability of calibration learning and avoids the ambiguity introduced by direct end-to-end correction.

3.5. Calibrated Output and Downstream Intelligent Perception

After the dual-branch network predicts the incremental parameters, the framework recursively updates the drift matrix and the mounting offset matrix according to

$$R_{DG} \leftarrow R_{DG}\Delta R_{DG}, \qquad R_{BS} \leftarrow \Delta R_{BS}R_{BS} \tag{17}$$

Here, the recursive update follows Eq. (17), where the drift state is updated by right multiplication and the mounting-offset state is updated by left multiplication. The updated matrices are then used to generate calibrated rotational and acceleration observations. As shown in Fig. 1, the calibrated outputs are denoted as R_{cal} and a_{cal}, which serve as corrected motion observations for downstream perception tasks.

The calibrated outputs can be directly integrated into higher-level application modules, including pose estimation, action recognition, movement analysis, and performance evaluation. The contribution of dynamic calibration therefore extends beyond the reduction of low-level sensor error to the improvement of robustness and reliability in downstream sports perception tasks. The current framework thus establishes a complete processing chain from wearable sensing and online calibration to downstream perception.

3.6. Online Recursive Calibration Pipeline

The bottom part of Fig. 1 illustrates the online recursive calibration pipeline more explicitly. The process starts with continuous IMU collection and buffer filling, followed by motion diversity evaluation and trigger decision. If updating is activated, the Transformer dual-branch parameter estimator predicts new calibration increments, which are then used to update R_{DG} and R_{BS}. The resulting calibrated observations are finally output and fed back into the loop for continuous recursive operation.

This design reflects the essential difference between the proposed framework and conventional offline calibration methods. Traditional calibration is performed only once before the experiment, assuming fixed sensor attachment and stable reference alignment. In contrast, the present method performs recursive updating throughout the motion process, thereby making the calibration state adaptive to time-varying dynamic sports conditions.

3.7. Training Objective

From a conceptual perspective, the proposed calibration framework aims to improve both observation-level consistency and parameter-level estimation accuracy. Accordingly, the overall calibration objective can be written as

$$J = w_1 L_{OME} + w_2 L_{AME} + w_3 L_{DG} + w_4 L_{BS}, \tag{18}$$

where L_{OME} denotes the orientation measurement error, L_{AME} denotes the acceleration measurement error, L_{DG} denotes the drift estimation error, and L_{BS} denotes the mounting-offset estimation error. Here, J represents the desired calibration objective from a system perspective and is used to describe the overall behavior expected from the framework.

In the current implementation, however, the supervisory signals are constructed from simulated drift and mounting-offset perturbations. Therefore, the network is optimized using parameter-level supervision on the two estimation branches rather than directly using observation-level metrics as gradient supervision. Specifically, the practical training loss is defined as

$$L_{\text{train}} = \lambda_{DG} \cdot \text{MSE}(\hat{d}^{DG}, d^{DG}) + \lambda_{BS} \cdot \text{MSE}(\hat{d}^{BS}, d^{BS}), \tag{19}$$

where $\hat{d}^{DG}$ and $\hat{d}^{BS}$ denote the predicted drift and mounting-offset parameters, while d^{DG} and d^{BS} denote the corresponding simulated targets. In the present experiments, equal weights are adopted for the two branches,

$$\lambda_{DG} = 1, \qquad \lambda_{BS} = 1. \tag{20}$$

Accordingly, the implemented training loss becomes

$$L_{\text{train}} = \text{MSE}(\hat{d}^{DG}, d^{DG}) + \text{MSE}(\hat{d}^{BS}, d^{BS}). \tag{21}$$

Under this formulation, OME and AME are not directly used for back-propagation during training; instead, they are used as observation-level evaluation metrics in the experiments to quantify how the learned calibration parameters improve corrected rotation and acceleration measurements. This design ensures consistency between the conceptual objective of the framework and the practical optimization procedure used in implementation.

By explicitly supervising the drift-estimation and mounting-offset-estimation branches during training, and by evaluating the resulting model from both observation-level and task-level perspectives, the proposed framework

Table 2. Statistics of the experimental datasets and sports-like proxy motion subsets.

Dataset / Subset	Subjects	Frames	Windows	Description
TIC test set	5	1,049,294	3,494	Original TIC evaluation data
High-dynamic sports-like proxy subset	5	366,900	1,223	Selected high-dynamic proxy windows
Explosive proxy	5	99,600	332	Jumping- and burst-like motion patterns
Arm-swing proxy	5	79,200	264	Punching- or racket-swing-like motion patterns
Turning proxy	5	77,400	258	Shuttle-run-, side-step-, and turning-like patterns
Coordination proxy	5	110,700	369	Squat-, lunge-, cross-step-, and aerobic-like combinations

improves calibration robustness and downstream perceptual reliability in high-dynamic sports-like motion scenarios.

4. Experiments and results

4.1. Experimental Setup

4.1.1. Dataset Description

The experimental dataset consisted of two components: the original TIC baseline dataset and a high-dynamic sports-like proxy subset. It should be noted that the second component is not an independently collected basketball, boxing, or sport-specific dataset. Instead, it is a proxy subset extracted from the TIC test sequences to approximate several high-dynamic motion characteristics that commonly appear in sports scenarios. The experimental data were divided according to the original TIC test-set protocol, where all available test sequences from the five subjects were used for baseline evaluation. The original TIC baseline dataset followed the training and test sets used in the TIC reproduction experiments[24] and served for baseline model training, evaluation, and comparison. To further evaluate the proposed method under sports-like high-dynamic conditions, a high-dynamic proxy subset was constructed from the TIC test set using a rule-based window selection strategy.

Specifically, each continuous test sequence was divided into non-overlapping windows of 300 frames. For each window, several motion-intensity indicators were calculated from the ground-truth or raw IMU signals, including acceleration variation, forearm angular velocity, head-pelvis rotation variation, and multi-IMU coordination intensity. These indicators were normalized and combined into a comprehensive dynamic score:

$$S = w_1 S_{acc} + w_2 S_{ang} + w_3 S_{rot} + w_4 S_{coord} \tag{22}$$

where S_{acc} measures acceleration changes, S_{ang} reflects rapid limb rotation, S_{rot} describes global or trunk orientation variation, and S_{coord} measures the coordination intensity across multiple IMUs. The weights were set uniformly unless otherwise specified, so that no single motion cue dominated the subset construction.

Windows with dynamic scores in the top 35% were selected as the high-dynamic sports-like proxy subset. According to the dominant motion characteristics, the selected windows were further grouped into four categories: explosive motion, arm-swing motion, turning or direction-changing motion, and continuous coordination motion. These category names describe motion-pattern proxies rather than manually annotated real sports actions. For example, the arm-swing and explosive groups are intended to approximate motion characteristics that may occur in boxing, racket sports, basketball jumping, or rapid sport maneuvers, but they should not be interpreted as direct samples from those sports. This rule-based construction avoids manual subjective labeling and ensures that the subset can be reproduced directly from the TIC test sequences.

This division strategy enables two levels of evaluation: the full TIC test set is used to assess overall calibration performance, while the high-dynamic proxy subset focuses on challenging motion segments that are more similar to sports scenarios, where rapid acceleration, large limb rotation, and accumulated calibration drift are more likely to occur. To provide a clearer description of the experimental data, the statistics of the original TIC test set and the high-dynamic sports-like proxy subset are summarized in Table 2.

4.1.2. Training Configuration and Model Hyperparameters

Following the experimental setup described above, the model was trained using data from six IMU sensors. The input dimension was set to 72. For the dual-branch output head, each branch produced a 36-dimensional vector, corresponding to the 6D rotation representations of six IMUs; therefore, the total output dimensionality of the network was 72. During training, the sequence length was fixed at 256 frames, while an online buffer length of 512 frames was used for inference. The raw IMU signals were sampled at 60 Hz and then processed into model inputs at 30 Hz.

The network was built with a Transformer Encoder architecture consisting of three encoder layers. The model dimension d_{model} was set to 256, with 8 attention heads in the multi-head attention module. The feed-forward dimension d_{ff} was set to 512. For optimization, the batch size

was set to 128, the learning rate was 1×10^{-3}, and Adam was used as the optimizer.

4.1.3. *Training Convergence Analysis*

To verify the stability of model training, the normalized training loss, drift error, and offset error are plotted in Fig. 2. All three curves decrease steadily during training, indicating that the model can effectively learn both global drift correction and local mounting-bias estimation. The convergence trend also supports the reliability of the subsequent calibration evaluation. The training loss decreases

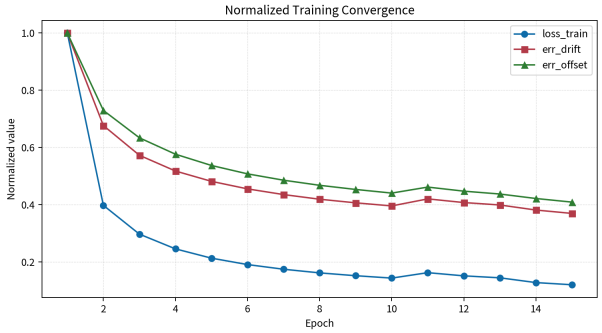

Fig. 2. Normalized training convergence of loss, drift error, and offset error

rapidly in the early epochs and then gradually stabilizes, while both drift error and offset error show a consistent downward trend. This suggests that the model does not only fit the training objective, but also improves the estimation of calibration-related rotation components. Therefore, the trained model provides a stable basis for the following static-dynamic comparison, sports-like proxy subset evaluation, and ablation experiments.

4.1.4. *Baseline Methods and Fair Comparison Settings*

To evaluate the proposed method, this study sets up several baseline and ablation experiments. E0: Static Only uses only the initial static calibration without online updates. E1: Original TIC represents the original dynamic calibration baseline. E2: Sports-TIC applies the original TIC method to the high-dynamic sports-like proxy subset to examine its performance under sport-relevant high-dynamic motion patterns.

To ensure fair comparison, all experiments use the same data split, preprocessing pipeline, input frequency, sequence construction strategy, and evaluation metrics whenever possible. The input IMU format, window length, and testing protocol are kept consistent across different experimental settings. Therefore, the observed performance differences mainly come from the calibration strategy rather than changes in data processing or evaluation conditions.

For ablation analysis, E3: w/o Trigger removes the action-diversity trigger, and E4: w/o Acc removes the acceleration-assisted branch. In addition, E5: Downstream Verification compares downstream task performance, such as pose estimation or action recognition, before and after calibration.

All comparison experiments follow the same data split, preprocessing procedure, and evaluation metrics whenever possible, so that the performance differences mainly reflect the effects of different calibration settings.

4.2. Evaluation Metrics

The following metrics are used for experimental evaluation. In the current implementation, OME and AME are treated as observation-level evaluation metrics rather than direct optimization objectives during training.

Let N denote the number of evaluated samples. For two rotation matrices $R_1, R_2 \in SO(3)$, the rotation error is measured by the geodesic distance on $SO(3)$:

$$d_R(R_1, R_2) = \arccos\left(\frac{\mathrm{tr}(R_1^\top R_2) - 1}{2}\right) \tag{23}$$

where $R_1^\top R_2$ is the relative rotation between R_1 and R_2, $\mathrm{tr}(\cdot)$ denotes the matrix trace, and $d_R(R_1, R_2)$ gives the angular distance in radians, which is reported in degrees in the result tables.

The first metric is Orientation Measurement Error (OME), which evaluates the correction quality of rotation observations after calibration:

$$\mathrm{OME} = \frac{1}{N}\sum_{i=1}^{N} d_R(\hat{R}_i^{\mathrm{corr}}, R_i^{\mathrm{gt}}) \tag{24}$$

where $\hat{R}_i^{\mathrm{corr}}$ is the corrected orientation matrix of the i-th sample after calibration, R_i^{gt} is the corresponding reference orientation matrix, and $d_R(\cdot,\cdot)$ is the geodesic rotation error defined in Eq. (23). Thus, OME measures the average orientation discrepancy between calibrated and reference rotations.

The second metric is Acceleration Measurement Error (AME), which measures the error between corrected acceleration and reference acceleration:

$$\mathrm{AME} = \frac{1}{N}\sum_{i=1}^{N} \left\| \hat{a}_i^{\mathrm{corr}} - a_i^{\mathrm{gt}} \right\|_2 \tag{25}$$

where $\hat{a}_i^{\mathrm{corr}} \in \mathbb{R}^3$ denotes the corrected acceleration vector of the i-th sample, $a_i^{\mathrm{gt}} \in \mathbb{R}^3$ denotes the corresponding reference acceleration vector, and $\|\cdot\|_2$ is the Euclidean norm. Therefore, AME measures the mean vector error between calibrated and reference accelerations.

Table 3. Overall comparison between static and dynamic calibration using Orientation Measurement Error (OME), Acceleration Measurement Error (AME), drift correction matrix error (R_{DG} Err), and mounting-offset correction matrix error (R_{BS} Err).

Method	OME	AME	R_{DG} Err	R_{BS} Err
E0: Static Only	49.9810°	2.5558 m/s^2	N/A	N/A
E1: Original TIC	15.2047°	1.3022 m/s^2	9.1776°	15.2804°

The third metric is drift correction matrix error (R_{DG} Err), which evaluates the estimation accuracy of the drift correction matrix:

$$R_{DG}\ \text{Err} = \frac{1}{N}\sum_{i=1}^{N} d_R(\hat{R}_{DG,i}, R^{\text{gt}}_{DG,i}) \tag{26}$$

where $\hat{R}_{DG,i}$ is the estimated drift rotation matrix for the i-th sample and $R^{\text{gt}}_{DG,i}$ is the corresponding reference drift matrix. This metric evaluates the mean geodesic error between estimated and reference global-drift rotations.

The fourth metric is mounting-offset correction matrix error (R_{BS} Err), which measures the estimation error of the sensor-to-body mounting-offset matrix:

$$R_{BS}\ \text{Err} = \frac{1}{N}\sum_{i=1}^{N} d_R(\hat{R}_{BS,i}, R^{\text{gt}}_{BS,i}) \tag{27}$$

where $\hat{R}_{BS,i}$ is the estimated sensor-to-body mounting-offset matrix for the i-th sample and $R^{\text{gt}}_{BS,i}$ is the corresponding reference mounting-offset matrix. This metric evaluates the mean geodesic error of local mounting-offset estimation.

For downstream verification, task-level metrics are introduced for the multi-class action recognition task. Classification Accuracy is used to evaluate the overall proportion of correctly classified samples, while Macro-F1 is adopted to provide a more balanced evaluation across different proxy categories. These downstream metrics are reported to examine whether improvements in calibration quality can be effectively transferred to downstream proxy action recognition performance under high-dynamic sports-like motion conditions.

4.3. Experimental Results

4.3.1. Comparison Between Static and Dynamic Calibration

This section evaluates whether online dynamic calibration outperforms one-time static calibration. Table 3 reports the overall comparison between E0: Static Only and E1: Original TIC using Orientation Measurement Error (OME), Acceleration Measurement Error (AME), drift correction matrix error (R_{DG} Err), and mounting-offset correction matrix error (R_{BS} Err).

The results indicate that dynamic calibration substantially improves performance relative to static calibration. Compared with E0, E1 reduces OME from 49.9810° to 15.2047°, corresponding to a reduction of 69.6%, and reduces AME from 2.5558 m/s^2 to 1.3022 m/s^2, corresponding to a reduction of 49.0%.

This improvement is mainly because static calibration only uses the initial calibration parameters, which may gradually become unreliable during high-dynamic sports-like motions. In contrast, dynamic calibration continuously estimates R_{DG} and R_{BS}, allowing the system to compensate for global drift and local mounting bias during motion. Since E0 does not estimate these two matrices online, the corresponding errors are marked as N/A.

Overall, online dynamic calibration is necessary for high-dynamic sports-like motion scenarios, as it provides more accurate orientation and acceleration measurements than one-time static calibration. The comparison in Fig. 3 further visualizes this improvement, showing that Original TIC consistently reduces both OME and AME compared with Static Only.

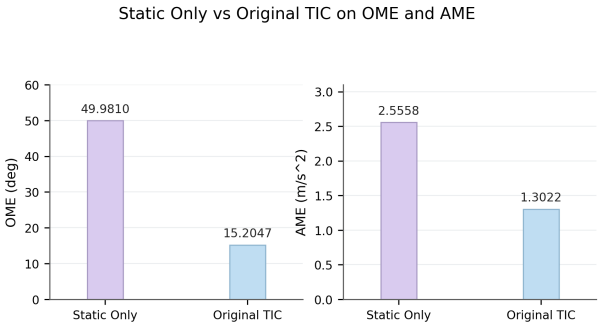

Fig. 3. Static Only vs Original TIC on Orientation Measurement Error (OME) and Acceleration Measurement Error (AME).

4.3.2. Performance on High-Dynamic Sports-Like Motion Subsets

This section evaluates whether the dynamic calibration method remains effective on high-dynamic sports-like proxy motions. Table 4 compares E1: Original TIC with E2: Sports-TIC across four proxy motion categories: explosive, arm swing, turning, and coordination. E1 serves as the reference result on the original TIC test set, whereas E2 reports performance on the high-dynamic sports-like

Table 4. Orientation Measurement Error (OME) and Acceleration Measurement Error (AME) comparison between E1 and E2 under different sports-like motion categories.

Action Category	E1: Original TIC OME	E1: Original TIC AME	E2: Sports-TIC OME	E2: Sports-TIC AME
Explosive	15.2047°	1.3022 m/s^2	15.6138°	2.0424 m/s^2
Arm Swing	15.2047°	1.3022 m/s^2	14.7528°	1.7021 m/s^2
Turning	15.2047°	1.3022 m/s^2	17.2239°	1.9510 m/s^2
Coordination	15.2047°	1.3022 m/s^2	13.9185°	1.8765 m/s^2

Table 5. Ablation study of trigger mechanism and acceleration assistance using Orientation Measurement Error (OME), Acceleration Measurement Error (AME), drift correction matrix error (R_{DG} Err), and mounting-offset correction matrix error (R_{BS} Err).

Method	OME	AME	R_{DG} Err	R_{BS} Err
E1: Original TIC	15.2047°	1.3022 m/s^2	9.1776°	15.2804°
E3: w/o Trigger	15.4836°	1.3046 m/s^2	9.2699°	15.5568°
E4: w/o Acc	63.5149°	9.2595 m/s^2	81.5929°	72.0373°

proxy subset.

Overall, Sports-TIC remains effective on the high-dynamic sports-like proxy subset. As shown in Table 4, the OME values of E2 remain close to the E1 baseline across different proxy categories, indicating that the dynamic calibration strategy preserves stable orientation correction under challenging motion conditions. To further visualize the category-wise differences, Fig. 4 presents the OME and AME comparison between E1 and E2 across the four proxy motion categories. For explosive actions, rapid acceleration fluctuations lead to relatively higher AME, whereas OME remains close to E1, indicating that the method maintains orientation correction accuracy. For arm-swing actions, upper-limb IMUs exhibit more pronounced local misalignment, making dynamic mounting-bias correction particularly beneficial. For turning actions, larger heading variations increase the importance of drift compensation. For coordination actions, continuous movement sequences are more susceptible to error accumulation, under which online recursive updating offers greater advantage.

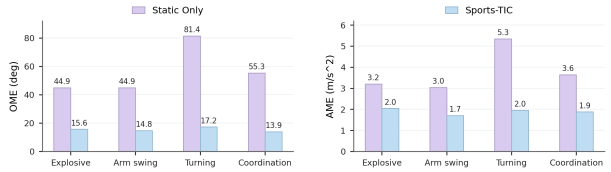

Fig. 4. Error comparison under different sports-like motion categories for drift correction matrix error (R_{DG} Err) and mounting-offset correction matrix error (R_{BS} Err).

These results indicate that the method remains effective in high-dynamic sports-like motion scenarios, particularly for actions characterized by large acceleration fluctuations, substantial heading changes, sensor-to-body misalignment, or prolonged continuous motion.

4.3.3. Ablation Study

This section evaluates the contribution of two key components in the dynamic calibration framework: the action-diversity trigger and acceleration assistance. Three settings are compared: E1: Original TIC, E3: w/o Trigger, and E4: w/o Acc. The results are shown in Table 5.

Compared with the complete model, E3: w/o Trigger yields slightly higher Orientation Measurement Error (OME), Acceleration Measurement Error (AME), drift correction matrix error (R_{DG} Err), and mounting-offset correction matrix error (R_{BS} Err). This finding indicates that not all motion windows are suitable for online updating and that the rotation-diversity trigger suppresses unreliable calibration updates.

The performance degradation is substantially more pronounced in E4: w/o Acc. Without acceleration input, all metrics increase markedly, indicating that acceleration provides critical complementary information for dynamic drift and mounting-bias estimation. This contribution is especially important in high-dynamic sports-like motions with large acceleration fluctuations.

The ablation results are visualized in Fig. 5.

4.3.4. Downstream Action Recognition Verification

To examine whether the calibration benefit can be transferred to downstream intelligent perception under different classifier settings, three action recognition models were evaluated: Logistic Regression, a multilayer perceptron (MLP), and a one-dimensional convolutional neural network (1D-CNN). Logistic Regression was used as a linear baseline. The MLP takes window-level IMU statistical fea-

Table 6. Downstream action recognition performance under subject-wise five-fold GroupKFold evaluation before and after dynamic calibration.

Classifier	Input Source	Accuracy	Macro-F1
Logistic Regression	Static IMU	0.351 ± 0.181	0.258 ± 0.114
Logistic Regression	TIC Calibrated IMU	0.423 ± 0.159	0.323 ± 0.097
MLP	Static IMU	0.420 ± 0.145	0.325 ± 0.090
MLP	TIC Calibrated IMU	0.425 ± 0.130	0.328 ± 0.069
1D-CNN	Static IMU	0.288 ± 0.211	0.188 ± 0.137
1D-CNN	TIC Calibrated IMU	0.449 ± 0.088	0.408 ± 0.070

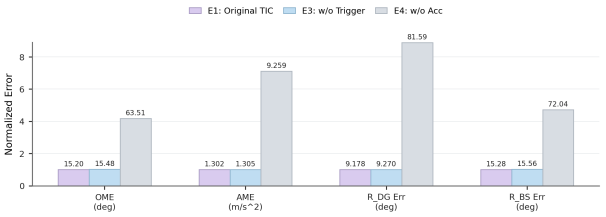

Fig. 5. Ablation comparison of E1, E3, and E4 across calibration metrics, including Orientation Measurement Error (OME), Acceleration Measurement Error (AME), drift correction matrix error (R_{DG} Err), and mounting-offset correction matrix error (R_{BS} Err).

tures as input to evaluate whether the improved feature distribution after calibration benefits a nonlinear classifier. In contrast, the 1D-CNN directly takes temporal signals composed of acceleration and 6D rotation representations from six IMUs as input, enabling the model to capture local temporal dynamic patterns during motion.

To avoid data leakage caused by samples from the same subject appearing in both training and testing sets, all classifiers were evaluated using subject-wise five-fold GroupKFold cross-validation. Accuracy and Macro-F1 were adopted as evaluation metrics. In addition, to examine whether the performance improvement of TIC dynamic calibration over static calibration was statistically consistent across subjects, paired t-tests were conducted on the fold-wise Accuracy and Macro-F1 scores, and Wilcoxon signed-rank tests were further reported as non-parametric supplements. The significance level was set to $p < 0.05$.

The results show that TIC-calibrated IMU inputs generally outperform Static IMU inputs across Logistic Regression, MLP, and 1D-CNN classifiers. The current downstream action labels are proxy labels constructed from motion characteristics in the TIC dataset, rather than manually annotated basketball, boxing, or other real sports action labels. Therefore, this experiment is used to evaluate whether improved calibration quality can promote downstream intelligent perception performance on sport-relevant motion proxies, rather than to establish a fully independent real sports action recognition benchmark.

As shown in Table 6, for Logistic Regression, Accuracy increases from 0.351 ± 0.181 to 0.423 ± 0.159, and Macro-F1 increases from 0.258 ± 0.114 to 0.323 ± 0.097. For MLP, Accuracy increases from 0.420 ± 0.145 to 0.425 ± 0.130, and Macro-F1 increases from 0.325 ± 0.090 to 0.328 ± 0.069. The most evident improvement is observed in the 1D-CNN model, which directly models IMU temporal dynamics: Accuracy increases from 0.288 ± 0.211 to 0.449 ± 0.088, and Macro-F1 increases from 0.188 ± 0.137 to 0.408 ± 0.070.

The statistical tests further show that the Macro-F1 improvement of 1D-CNN reaches statistical significance under the paired t-test ($p = 0.0211$), while the Accuracy improvement shows a near-significant trend ($p = 0.0655$). These results indicate that the proposed TIC dynamic calibration method not only reduces IMU orientation and acceleration errors, but also transfers the improvement in sensor observation quality to downstream temporal action recognition tasks.

Fig. 6 visualizes the classifier-wise comparison between Static IMU and TIC-calibrated IMU inputs for Accuracy and Macro-F1, with statistical significance marked for the 1D-CNN Macro-F1 improvement.

Overall, these results suggest that online dynamic calibration not only improves low-level sensor observation quality but also enhances downstream proxy action recognition performance under high-dynamic sports-like motion conditions, especially when temporal dynamics are explicitly modeled by a deep learning classifier.

4.4. Results Analysis

4.4.1. Analysis of Online Dynamic Calibration Effect

The results in Section 4.3.1 indicate that online dynamic calibration provides substantial improvements over one-time static calibration in high-dynamic sports-like motion scenarios. Compared with E0: Static Only, E1: Original TIC significantly reduces both OME and AME, indicating that static calibration parameters become progressively unreliable during long-duration and high-intensity movements, whereas the proposed online strategy continuously

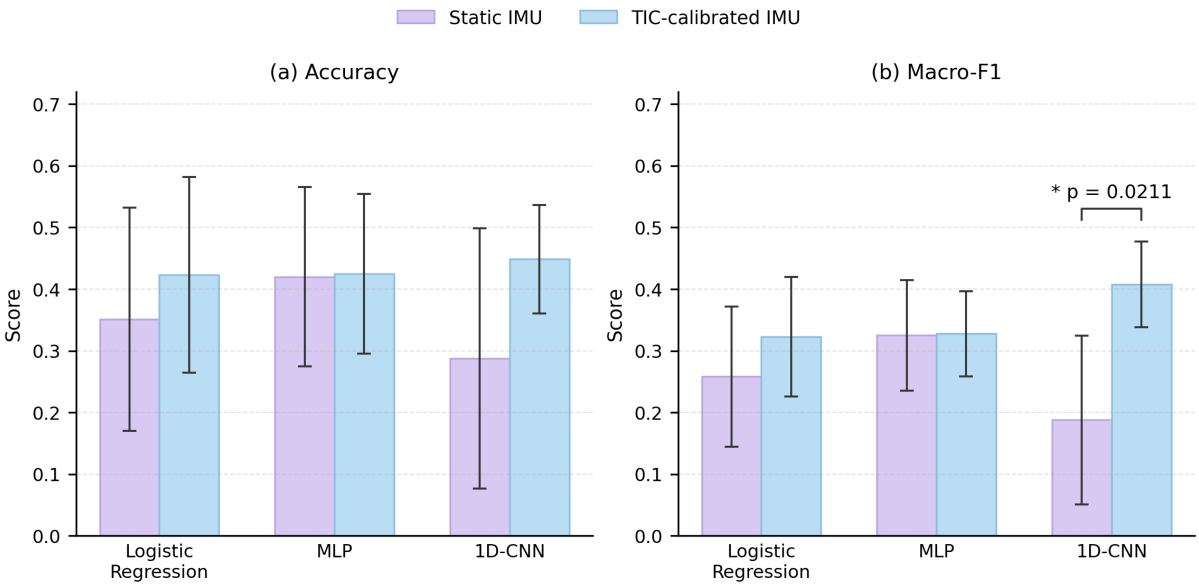

Fig. 6. Comparison of downstream action recognition performance across Logistic Regression, MLP, and 1D-CNN using Static IMU and TIC-calibrated IMU inputs. Error bars denote standard deviations across subject-wise five-fold GroupKFold evaluation.

corrects sensor observations during motion. This result confirms that, in high-dynamic sports-like actions, calibration should be treated not as a one-time initialization step, but as a time-varying process requiring continuous compensation.

The performance on the high-dynamic sports-like proxy subset further supports this conclusion. Although different proxy categories exhibit different motion characteristics, the proposed dynamic calibration strategy maintains stable orientation correction across explosive, arm-swing, turning, and coordination proxy motions. In particular, the relatively small variation in OME across the four proxy motion categories indicates that the framework can preserve rotational correction quality even under severe posture changes and rapid movement transitions. Meanwhile, the changes in AME across categories suggest that acceleration correction is more sensitive to action intensity and movement pattern, especially in explosive and turning proxies. Overall, these findings demonstrate that the proposed method is not only effective on the original benchmark setting, but also promising for more challenging high-dynamic sports-like motion scenarios, while direct validation on real sport-specific data remains necessary.

4.4.2. Analysis of Motion-Diversity Trigger Effect

The ablation results in Section 4.3.3 reveal that the motion-diversity trigger contributes to the stability of online dynamic calibration. When this trigger is removed, performance degrades slightly but consistently across Orientation Measurement Error (OME), Acceleration Measurement Error (AME), drift correction matrix error (R_{DG} Err), and mounting-offset correction matrix error (R_{BS} Err). This observation indicates that not every temporal window is equally suitable for calibration updating. In low-motion or weak-information segments, forcing the framework to update calibration parameters may introduce unstable corrections or unnecessary fluctuations. By contrast, the motion-diversity trigger allows the system to selectively update calibration only when the current motion contains sufficient informative variation, thereby improving update reliability.

The comparison with E4: w/o Acc further highlights the importance of informative motion cues in dynamic calibration. Without acceleration input, all error metrics increase drastically, which indicates that acceleration provides critical complementary information for estimating both global drift and local mounting bias. This effect is particularly important in high-dynamic sports-like motions, where rapid changes in body motion cannot be adequately characterized by rotational information alone. Therefore, the complete framework benefits from a two-level stabilizing mechanism: the motion-diversity trigger determines when the calibration should be updated, while the acceleration-assisted branch improves how the calibration parameters are estimated. This explains why the full model achieves the most balanced performance among all compared set-

tings.

4.4.3. Analysis of Perception-Oriented Practical Value

Beyond reducing low-level sensor errors, the practical value of the proposed framework lies in its contribution to downstream proxy action recognition. The downstream verification results show that TIC-calibrated IMU inputs generally outperform Static IMU inputs across Logistic Regression, MLP, and 1D-CNN classifiers. In particular, the 1D-CNN model, which directly models IMU temporal dynamics, shows the most evident improvement: Accuracy increases from 0.288 ± 0.211 to 0.449 ± 0.088, and Macro-F1 increases from 0.188 ± 0.137 to 0.408 ± 0.070. The paired t-test further indicates that the Macro-F1 improvement of 1D-CNN reaches statistical significance ($p = 0.0211$). These improvements indicate that the calibrated IMU observations provide more reliable and more discriminative motion information for downstream proxy action recognition. In other words, the benefit of dynamic calibration is not limited to sensor-level consistency, but can be effectively transferred to task-level classification performance under proxy motion labels.

This perception-oriented gain is particularly meaningful for sports-like motion analysis applications. In high-dynamic sports-like motion scenarios, downstream motion classification is highly sensitive to cumulative sensor drift, local mounting deviation, and unstable motion representation. By improving the quality of rotation and acceleration observations at the input stage, the proposed framework enhances the robustness of proxy motion category discrimination. Although a performance gap still remains between the calibrated IMU input and the ground-truth reference, the current results suggest that online dynamic calibration can serve as a useful front-end component for high-dynamic sports-like motion perception systems. Future work may further strengthen this effect by combining the proposed calibration framework with more advanced downstream action recognition backbones, class-specific temporal modeling strategies, and real sport-specific annotations.

4.4.4. Experimental Limitations and Future Work

Although the proposed dynamic calibration method effectively reduces IMU orientation and acceleration errors, several limitations remain. First, the experiments are mainly based on the TIC Dataset. The constructed high-dynamic sports-like proxy subset improves scenario relevance, but it is still derived from general human motion data rather than real competitive sports. Therefore, it does not fully represent sport-specific dynamics such as basketball jumps and direction changes, boxing strikes, body contact, high-frequency impacts, and severe sensor slippage. Its motion intensity, impact frequency, and wearing-condition variability may consequently be lower than those in real basketball, football, badminton, boxing, or combat training scenarios.

Second, the high-dynamic proxy subset is selected using rule-based indicators, including acceleration variation, angular velocity, and multi-IMU coordination. This improves reproducibility, but may not fully cover complex sports-specific patterns such as sudden stops, direction changes, jumping and landing, punching impact, body contact, and continuous racket swings.

Third, the current trigger mechanism relies on fixed motion diversity thresholds τ_{DG} and τ_{BS}. Although this suppresses unreliable updates in low-information windows, fixed thresholds may be less adaptive to different sports, athletes, and sensor attachment conditions. In real sports scenes, rapidly changing motion intensity may cause over-triggering or insufficient triggering.

Future work will focus on collecting real sport-specific IMU datasets with high-precision reference systems, especially for basketball, boxing, racket sports, and combat training. Such datasets should cover actions such as rapid turning, jumping, landing, striking, body contact, high-frequency impact, and long-duration continuous motion, so that the method can be validated beyond proxy motion subsets. The trigger strategy can also be improved by adaptively adjusting τ_{DG} and τ_{BS} according to motion intensity, signal confidence, noise level, and historical update stability. In addition, incorporating impact detection, motion-phase recognition, uncertainty estimation, and inter-IMU consistency constraints may further improve robustness against strap slippage, soft-tissue artifacts, and extreme dynamic disturbances. For real-time deployment, lightweight Transformer structures or model compression methods can be explored to reduce inference latency while maintaining calibration accuracy.

5. Conclusion

This paper presents an adaptive IMU dynamic calibration framework for high-dynamic sports-like motion perception. By combining a motion-diversity-aware trigger with a Transformer Encoder-based dual-branch calibration network, the proposed method enables online recursive estimation of sensor drift and mounting offset under highly dynamic motion conditions. Experiments on the TIC dataset and the constructed high-dynamic sports-like proxy subset show that the proposed method effectively reduces orientation and acceleration errors. Downstream proxy action recognition results further indicate that dynamically calibrated IMU observations improve both Accuracy and

Macro-F1 relative to static IMU inputs, demonstrating that calibration gains can be transferred from sensor-level correction to task-level classification performance under proxy motion labels.

Overall, the proposed method provides a practical solution for online IMU calibration in high-dynamic motion environments and highlights the importance of adaptive calibration for reliable motion perception systems.

Acknowledgment

This research was supported by the Heilongjiang Provincial Philosophy and Social Science Research Planning Project (Project No. 25TYE010).

References

[1] A. Ç. Seçkin, B. Ateş, and M. Seçkin, (2023) *"Review on wearable technology in sports: concepts, challenges and opportunities"* **Applied sciences *13***(18): 10399. DOI: 10.3390/app131810399.

[2] P. Picerno, M. Iosa, C. D'Souza, M. G. Benedetti, S. Paolucci, and G. Morone, (2021) *"Wearable inertial sensors for human movement analysis: a five-year update"* **Expert review of medical devices *18***(sup1): 79–94. DOI: 10.1080/17434440.2021.1988849.

[3] W. Hailong, G. Xinru, J. Sheng, and Z. Lijun, (2026) *"Multi-Task Learning for Sports Training Monitoring Via Multimodal Fusion of IMU and Human-Body Capacitance Signals"* **Journal of Applied Science and Engineering *31***: 1–13. DOI: 10.6180/jase.202608_31.038.

[4] M. Ekdahl, A. Loewen, A. Erdman, S. Sahin, and S. Ulman, (2023) *"Inertial measurement unit sensor-to-segment calibration comparison for sport-specific motion analysis"* **Sensors *23***(18): 7987. DOI: 10.3390/s23187987.

[5] Y.-t. Zhao, X. Wang, K.-x. Yang, L.-d. Wang, C. Yang, R. Feng, L.-l. Zheng, and Z.-p. Zhou, (2025) *"Effects of various static calibration postures on knee mechanics during locomotor tasks using statistical parametric mapping analysis"* **Scientific Reports *15***(1): 29833. DOI: 10.1038/s41598-025-15311-2.

[6] J. Wu, Z. Mo, X. Gao, W. Xin, W. Shi, and J. Park, (2025) *"Artificial intelligence assisted wearable flexible sensors for sports: research progress in technology integration and application"* **International Journal of Smart and Nano Materials *16***(3): 510–548. DOI: 10.1080/19475411.2025.2519582.

[7] J. Ok, S. Park, Y. H. Jung, and T.-i. Kim, (2024) *"Wearable and implantable cortisol-sensing electronics for stress monitoring"* **Advanced Materials *36***(1): 2211595. DOI: 10.1002/adma.202211595.

[8] X. Xuan, C. Chen, A. Molinero-Fernandez, E. Ekelund, D. Cardinale, M. Swarén, L. Wedholm, M. Cuartero, and G. A. Crespo, (2023) *"Fully integrated wearable device for continuous sweat lactate monitoring in sports"* **ACS sensors *8***(6): 2401–2409. DOI: 10.1021/acssensors.3c00708.

[9] F. Criscuolo, I. N. Hanitra, S. Aiassa, I. Taurino, N. Oliva, S. Carrara, and G. De Micheli, (2021) *"Wearable multifunctional sweat-sensing system for efficient healthcare monitoring"* **Sensors and Actuators B: Chemical *328***: 129017. DOI: 10.1016/j.snb.2020.129017.

[10] A. Bonfiglio, D. Tacconi, R. M. Bongers, and E. Farella, (2024) *"Effects of IMU sensor-to-segment calibration on clinical 3D elbow joint angles estimation"* **Frontiers in Bioengineering and Biotechnology *12***: 1385750. DOI: 10.3389/fbioe.2024.1385750.

[11] S. Chidambaram, Y. Maheswaran, K. Patel, V. Sounderajah, D. A. Hashimoto, K. P. Seastedt, A. H. McGregor, S. R. Markar, and A. Darzi, (2022) *"Using artificial intelligence-enhanced sensing and wearable technology in sports medicine and performance optimisation"* **Sensors *22***(18): 6920. DOI: 10.3390/s22186920.

[12] C. Yi, S. Zhang, F. Jiang, J. Liu, Z. Ding, C. Yang, and H. Zhou, (2021) *"Enable fully customized assistance: A novel IMU-based motor intent decoding scheme"* **IEEE Transactions on Cognitive and Developmental Systems *15***(4): 2089–2098. DOI: 10.1109/tcds.2021.3126001.

[13] T. Li and H. Yu, (2023) *"Upper body pose estimation using a visual–inertial sensor system with automatic sensor-to-segment calibration"* **IEEE Sensors Journal *23***(6): 6292–6302. DOI: 10.1109/jsen.2023.3241084.

[14] J. Ghattas and D. N. Jarvis, (2024) *"Validity of inertial measurement units for tracking human motion: a systematic review"* **Sports biomechanics *23***(11): 1853–1866. DOI: 10.1080/14763141.2021.1990383.

[15] J. Zhu, Z. Ye, R. Liu, and J. Liu, (2026) *"Inertial measurement units (IMUs) for biomechanical analysis in sport: A review of applications, challenges and future directions"* **Sensor Review *46***(1): 88–104. DOI: 10.1108/SR-04-2025-0261.

[16] M. McInne, D. Blana, A. Starkey, and E. K. Chadwick, (2025) *"A practical sensor-to-segment calibration method for upper limb inertial motion capture in a clinical setting"* **IEEE Journal of Translational Engineering in Health and Medicine**: DOI: 10.1109/JTEHM.2025.3565986.

[17] L. Wolski, M. Halaki, C. E. Hiller, E. Pappas, and A. Fong Yan, (2024) *"Validity of an inertial measurement unit system to measure lower limb kinematics at point of contact during incremental high-speed running"* **Sensors *24***(17): 5718. DOI: 10.3390/s24175718.

[18] M. V. Potter, (2025) *"Simulating effects of sensor-to-segment alignment errors on IMU-based estimates of lower limb joint angles during running"* **Sports Engineering *28***(1): 1. DOI: 10.1007/s12283-024-00483-3.

[19] B. Fan, L. Zhang, S. Cai, M. Du, T. Liu, Q. Li, and P. Shull, (2025) *"Influence of sampling rate on wearable IMU orientation estimation accuracy for human movement analysis"* **Sensors *25***(7): 1976. DOI: 10.3390/s25071976.

[20] S. Suh, V. F. Rey, and P. Lukowicz, (2023) *"Tasked: transformer-based adversarial learning for human activity recognition using wearable sensors via self-knowledge distillation"* **Knowledge-Based Systems *260***: 110143. DOI: 10.1016/j.knosys.2022.110143.

[21] X. Guo, Y. Kim, X. Ning, and S. D. Min, (2025) *"Enhancing the transformer model with a convolutional feature extractor block and vector-based relative position embedding for human activity recognition"* **Sensors *25***(2): 301. DOI: 10.3390/s25020301.

[22] D. Kim, Y. Jin, H. Cho, T. Jones, Y. M. Zhou, A. Fadaie, D. Popov, K. Swaminathan, and C. J. Walsh, (2025) *"Learning-based 3D human kinematics estimation using behavioral constraints from activity classification"* **Nature communications *16***(1): 3454. DOI: 10.1038/s41467-025-58624-6.

[23] D. Fruet, A. Tauro, D. Di Liberto, C. Pedrotti, I. Bracci, G. Martinelli, and G. Nollo. "Comparing IMU and Optical Motion Capture System for Sport Biomechanics: Static and Dynamic Analysis". In: *2025 IEEE International Workshop on Sport, Technology and Research (STAR)*. IEEE. 2025, 43–48. DOI: 10.1109/star66750.2025.11264765.

[24] C. Zuo, J. Huang, X. Jiang, Y. Yao, X. Shi, R. Cao, X. Yi, F. Xu, S. Guo, and Y. Qin, (2025) *"Transformer IMU calibrator: dynamic on-body IMU calibration for inertial motion capture"* **ACM Transactions on Graphics (TOG) *44***(4): 1–14. DOI: 10.1145/3730937.