

Article

Motion Sequence Analysis Using Adaptive Coding with Ensemble Hidden Markov Models

Xiangzeng Kong ^{1,†}, Xinyue Liu ^{2,†} , Shimiao Chen ², Wenxuan Kang ³, Zhicong Luo ¹, Jianjun Chen ⁴  and Tao Wu ^{2,*}

¹ College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou 350100, China; xzkong@fafu.edu.cn (X.K.)

² School of Future Technology, Fujian Agriculture and Forestry University, Fuzhou 350002, China

³ College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China

⁴ Department of Computing, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China

* Correspondence: wtscs@163.com

† These authors contributed equally to this work.

Abstract: Motion sequence data comprise a chronologically organized recording of a series of movements or actions carried out by a human being. Motion patterns found in such data holds significance for research and applications across multiple fields. In recent years, various feature representation techniques have been proposed to carry out sequence analysis. However, many of these methods have not fully uncovered the correlations between elements in sequences nor the internal interrelated structures among different dimensions, which are crucial to the recognition of motion patterns. This study proposes a novel Adaptive Sequence Coding (ASC) feature representation with ensemble hidden Markov models for motion sequence analysis. The ASC adopts the dual symbolization integrating first-order differential symbolization and event sequence encoding to effectively represent individual motion sequences. Subsequently, an adaptive boost algorithm based on a hidden Markov model is presented to distinguish the coded sequence data into different motion patterns. The experimental results on several publicly available datasets demonstrate that the proposed methodology outperforms other competing techniques. Meanwhile, ablation studies conducted on ASC and the adaptive boost approach further verify their significant potential in motion sequence analysis.

Keywords: motion sequence; dual symbolization; event encoding; hidden Markov model; adaptive boost

MSC: 68T01; 92C55



Citation: Kong, X.; Liu, X.; Chen, S.; Kang, W.; Luo, Z.; Chen, J.; Wu, T. Motion Sequence Analysis Using Adaptive Coding with Ensemble Hidden Markov Models. *Mathematics* **2024**, *12*, 185. <https://doi.org/10.3390/math12020185>

Academic Editor: Alexander Zeifman

Received: 26 November 2023

Revised: 18 December 2023

Accepted: 25 December 2023

Published: 5 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Motion sequence data comprise a kind of multi-dimensional time series that describes a motion phenomenon from various dimensions over time [1–3]. Motion sequence data are commonly found in real-world activities such as sports [4,5], traffic transportation prediction [6], smart home security threat detection [7], and medical rehabilitation [8]. Using aerobics training as an example, the motion sequence data collected from athletes during their training activities can be analyzed to better guide their strategy for difficult movements, thereby enhancing the effectiveness of training [9]. The intelligent analysis of motion sequences has also attracted widespread attention from the healthcare research field in recent years [2,8]. However, motion sequence data exhibit characteristics such as high dimensionality and strong coupling, and they are highly susceptible to noise caused by environmental interference [10,11]; all of these factors pose difficulties for the detection and analysis of motion signals manually. Therefore, a reliable approach should be developed to accurately and autonomously process complicated motion data.

Raw motion sequence data should first be transformed into structured and meaningful representations using feature representation methods to reduce data dimension and computational cost [12–14]. In addition, this process improves the signal-to-noise ratio, enabling traditional classification approaches to more accurately analyze the transformed sequences [10,15].

Existing representation methods can be roughly categorized into two groups: non-data-adaptive and data-adaptive methods [16,17]. Non-data-adaptive methods rely on predefined transformation parameters or structures for handling input data. They cannot be dynamically adjusted to adapt to the characteristics of a specific dataset during processing. In contrast, data-adaptive methods are capable of dynamically adjusting internal representations to learn and capture intricate patterns and variations present in the data. Due to their flexible representational capacity and excellent generalization performance, data-adaptive methods can reduce the need for manual intervention [18] and have been widely adopted to handle various types of multi-dimensional sequence data [19]. Common data-adaptive approaches include Symbolic Aggregation approximation (SAX) and its variants [18,20]. For instance, the work of Lima et al. [21] revealed that symbolic representation techniques (e.g., SAX, Symbolic Fourier Aggregation (SFA)) can reduce the computational cost of human activity recognition. Nguyen et al. [22] presented a framework based on symbolic representation and the linear classification model to address a series of challenges including the adaptation of variable-length motion data and interpretability.

Once all of the features have been identified for a motion sequence, an appropriate classification approach needs to be designed to discriminate between the motion patterns. Widely used classification approaches include but are not limited to support vector machine [21], hidden Markov model (HMM) [23], Naive Bayes [24], K-nearest neighbor (kNN) [25], and adaptive boosting (AdaBoost) [26]. Among these methods, AdaBoost has been extensively applied to classify motion sequences [26,27] as it can adaptively adjust the weights of base learners and sequence samples via a learning process, thereby improving the performance of sequence analysis.

Despite the remarkable success achieved by symbolization-based sequence analysis, especially those rooted in data-adaptive methods, some limitations still exist that researchers should pay attention to. First, most symbolization-based feature representations ignore intrinsic structure relationships in different dimensions of sequences and the correlations between elements in sequences, which are crucial clues for identifying motion patterns. Moreover, since some existing symbolization strategies are very sensitive to outliers or noise in the sequence data, certain information unrelated to classification may be identified, adversely affecting performance. Second, the classification accuracy of single classifiers may be constrained, especially in small sample size cases, since their dependence on a particular set of parameters hampers generalization across diverse datasets or distinct feature sets with disparate distributions.

To address the above challenges, we propose a novel motion sequence analysis framework that adaptively encodes and compresses multi-dimensional motion sequences via dual symbolization and assembles multiple base learners for efficient and accurate classification. More specifically, this framework first applies a first-order differential symbolization to find the correlations between elements and denoise the raw sequences. Then, a new adaptive event coding is devised to encode multi-dimensional symbol sequences into one-dimensional event sequences, with the objective of uncovering the inherent relationships within motion sequences across different dimensions. Finally, an ensemble learning classifier AdaBoost using HMM as the base learner is applied to discriminate between event sequences. The framework is used to carry out experiments on a series of real-life motion sequence datasets, and the experiment results evidence its superiority over its competitors. In addition, ablation experiments demonstrate the effectiveness of the devised feature representation mechanism in dimensional reduction.

Overall, this study has the following major contributions:

1. We propose a novel feature representation approach, namely, adaptive sequence coding (ASC), for motion data analysis. ASC is a data-adaptive learning method that does not require large-scale hyperparameters, which has the ability to capture the correlations between elements as well as the internal interrelated structures of multi-dimensional sequences.
2. We propose an ensemble learning classifier that utilizes HMMs as base learners. It can effectively excavate internal interconnections and variations of elements within symbolized sequences, thus further boosting the performance of motion recognition.
3. Extensive experiments on several popular real-world datasets show that our method compares well to competing techniques. Additionally, ablation studies also confirm the benefits of the proposed dual symbolization mechanism and ensemble learning.

The rest of this paper is organized as follows. Section 2 provides a literature review of related methods. Section 3 describes, in detail, the datasets used and the principle of our proposed method. Section 4 provides an analytical discussion of the experimental results. Finally, we give our conclusions and outline recommendations for future work in Section 5.

2. Related Works

A significant number of studies have been conducted for the identification and analysis of motion sequence data in recent years. They can be roughly grouped into two groups: no-data-adaptive and data-adaptive approaches. In this section, we briefly review these two categories of approaches.

2.1. Non-Data-Adaptive Methods

Non-data-adaptive methods involve analyzing and predicting data based on pre-determined rules and parameters without any kind of data adjustment or training. Spectrum analysis is a common non-data-adaptive representation approach [28,29]. For example, Agrawal et al. [28] proposed a feature representation approach based on the Discrete Fourier Transform, which maps time series data to the frequency domain and then performs similarity queries using R*-trees. Liu et al. [30] utilized the Discrete Wavelet Transform to jointly represent the time-frequency domain information of data. In addition, some other non-data-adaptive methodologies in [29,31,32] have also been proposed for the analysis of motion sequence. Keogh et al. [33] proposed a method based on segment-by-segment linear segmentation for determining the shape of the sequences. Similarly, Chatzigeorgakidis et al. [31] adopted a piecewise aggregate approximation to extract the representative features from the sequence data, which divides the original data into segments and then calculates the average value of each segment as a representative feature. Overall, non-data-adaptive methods are fast and simple due to their minimal reliance on training data and model parameter adjustments, but their strong dependence on a priori assumptions can often lead to underfitting or high error classification results, especially when dealing with complex motion sequences [29].

2.2. Data-Adaptive Methods

Data-adaptive methods have garnered significant attention in the analysis of sequence data, due to the fact that they can adaptively change transformation parameters when transforming sequences. In this regard, Lin et al. [18] first introduced a SAX method for the feature representation of sequence data. It utilizes the piecewise aggregate approximation to process the original sequence and partition the distribution space of the sequences into equiprobable regions according to the standard normal distribution. Then, each segment is mapped to the corresponding symbols. The SAX can preserve the overall shape and local features of the original sequence, making the distance between the characterized sequence and the original one strongly correlated. Due to its excellent characteristics, many SAX-based approaches have been proposed to analyze motion sequences involving humans and objects in the literature [21,34]. For instance, Junejo et al. [35] proposed a human motion recognition method based on profiles and SAX-Shapes, which transforms the raw

data into symbol sequences by shape information to reduce dimensionality while retaining important information. Recently, Zhou et al. [36] applied the SAX to explore the relationship between gaze focus and hand movement. Later, Riza et al. [37] used the SAX and Basis Order Discovery Random Projection algorithm to detect orbital resonance data describing asteroid motion. However, most existing symbolization-based motion sequence approaches ignore noise reduction processing and correlation information in various dimensions of sequences and among different elements in each one-dimensional sequence, negatively affecting classification performance. Thus, this work focuses on proposing an innovative feature representation mechanism to effectively tackle these challenges.

3. Materials and Methods

This section presents a novel framework for enhancing the analysis performance of motion sequences. The machine learning pipeline used in this research is depicted in Figure 1. It mainly encompasses two phases: feature representation learning and classification. Detailed discussions of each phase are presented in the subsequent subsections.

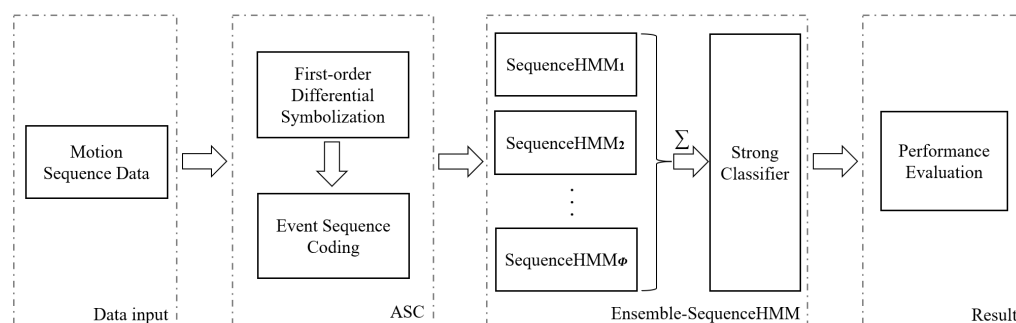


Figure 1. Flowchart of the proposed motion sequence analysis framework.

3.1. Dataset

Five publicly available datasets on human motion are used in this research. Their details are displayed in Table 1. Among these datasets, LIBRAS1, LIBRAS2, HAR, and JSI are relatively small and OPPORTUNITY has a larger sample size. It should also be noted that JSI and OPPORTUNITY are unbalanced datasets, while the other three datasets are balanced.

The first two datasets, LIBRAS1 and LIBRAS2, are derived from the same public dataset called Lingua Brasileira de Sinais (LIBRAS) [38], which is commonly used for Brazilian gesture language recognition studies. Both datasets comprise various instances of gesture movements, with each instance represented by 45 two-dimensional data trajectories. In LIBRAS1, the gestures include swing (curved), swing (horizontal), swing (vertical), arc (anti-clockwise), and arc (clockwise); in LIBRAS2, the gestures include zigzag (vertical), wavy (horizontal), wavy (vertical), circle, and curve (up).

The third dataset is a Human Activity Recognition (HAR) dataset [39]. It was collected from 30 participants aged from 19 to 48 years old using smartphones attached to the left side and right side of their waists. The participants performed daily life activities such as sitting and walking. The corresponding triaxial accelerometer (linear acceleration) and gyroscope (angular velocity) data were collected at a sampling rate of 50 Hz.

The fourth dataset selected is the localization data provided by the Jozef Stefan Institute (JSI) [40]. This database contains recordings of five individuals performing various activities. Each wore four sensors (tags) to record the motion data while repeating the same scenario five times. The 427 samples in this dataset were selected, and each sample is associated with one of four human motions (e.g., sitting down, walking).

The fifth dataset, known as the OPPORTUNITY Activity Recognition dataset [3], is introduced to serve as a benchmark for algorithms aimed at recognizing human motion recorded by a multitude of body-worn sensors. The dataset incorporates numerous inertial sensors embedded in everyday objects, along with tags and switches distributed throughout

the environment. The dataset comprises 1672 samples, and each sample contains a three-dimensional motion sequence capturing human motion, covering four primary activity types: standing, walking, sitting, and lying down.

Table 1. Detailed information of the experimental datasets

Datasets	Classes	Dimensions	Sequence Lengths	Sample Sizes of Sequences
LIBRAS1	5	2	[45, 45]	120
LIBRAS2	5	2	[45, 45]	120
HAR	4	3	[10, 48]	242
JSI	4	3	[133, 133]	427
OPPORTUNITY	4	3	[4028, 4028]	1672

3.2. Adaptive Motion Sequence Coding

In general, motion sequence data has high dimensionality and strong coupling. Sequence representation techniques can transform the original multi-dimensional sequences into an approximate representation, improving the detection performance of subsequent tasks. Among these methods, symbolic representation has emerged as a research hotspot for sequence representation learning. As mentioned above, SAX utilizes time-domain information to extract the important features, and SFA processes sequence data in the frequency domain. However, they only symbolize the sequence data on a single dimension and may lose important correlation information between sequences. To solve this issue, this work presents a novel dual symbolization framework to capture both time-wise and dimension-wise information of the motion sequence data. We begin by introducing the notation used throughout the paper. In what follows, the motion sequence dataset is denoted by $SDB = \{(X^{(1)}, y_1), (X^{(2)}, y_2), \dots, (X^{(n)}, y_n), \dots, (X^{(N)}, y_N)\}$, with $X^{(n)} = (X_1^{(n)}; X_2^{(n)}; \dots; X_d^{(n)}; \dots; X_D^{(n)})$, $n = 1, 2, \dots, N$; and $d = 1, 2, \dots, D$. Here, N denotes the total number of motion sequence samples, D denotes the dimensions of the sequence data, $y_n \in \{1, 2, \dots, K\}$ represents the class label of the n -th sequence sample, and K represents the total number of motion categories in the dataset. Note that D is usually 2 or 3 in this work. The d -th dimension data of the motion sequence $X^{(n)}$ can be represented as $X_d^{(n)} = \{x_{d1}^{(n)}, x_{d2}^{(n)}, \dots, x_{dt}^{(n)}, \dots, x_{dT}^{(n)}\}$, where T and $x_{dt}^{(n)} \in \mathbb{R}$ denote the sequence length and the t -th observation ($t = 1, 2, \dots, T$), respectively.

Since motion sequences are highly susceptible to noise caused by various sources such as sensor errors and environmental effects, a first-order difference method is adopted to smooth out the data and highlight the changes or trends between consecutive observations. The logic of this method can be mathematically expressed as [14]:

$$\nabla x_{dt}^{(n)} = x_{dt}^{(n)} - x_{d(t-1)}^{(n)}, \quad t = 2, 3, \dots, T. \quad (1)$$

where $\nabla x_{dt}^{(n)}$ represents the first-order difference at time t . Subsequently, a novel dual symbolization mechanism is designed to analyze the coupling relationship between multi-dimensional sequences and reduce the effect of the noise and outliers. Similar to SAX, the newly presented mechanism first transforms the difference data into a one-dimensional symbolic sequence quickly, using the following processing rule:

$$s_{dt}^{(n)} = \begin{cases} A, & \nabla x_{dt}^{(n)} > \varepsilon \\ B, & |\nabla x_{dt}^{(n)}| \leq \varepsilon \\ C, & \nabla x_{dt}^{(n)} < -\varepsilon \end{cases} \quad (2)$$

where $s_{dt}^{(n)}$ denotes the symbolic representation of the value at time t in difference data ($t = 2, 3, \dots, T$) and ε is a small positive real value that represents the sensitivity to spatial variation. Based on this, the first-order difference sequence can be represented using the

following sequence of symbols $S_d^{(n)} = \{s_{d2}^{(n)}, s_{d3}^{(n)}, \dots, s_{dt}^{(n)}, \dots, s_{dT}^{(n)}\}$. Symbols “A”, “B”, and “C” are the three potential trend changes calculated from the value of the sequence in the current moment with respect to the value from the previous moment. To be more specific, the positive and negative values are represented by the symbols “A” and “C”, respectively; when the absolute value of the difference is less than ε , it means that the value at successive moments is almost unchanged, which is denoted by the symbol “B”. In this way, the newly presented first-order difference symbolization strategy would capture the correlations between elements in each unidimensional series, such as variation trends.

Although symbolized sequences from the multi-dimensional data can be directly flattened into a one-dimensional feature vector as the input of classifiers, this might result in the loss of correlation information among distinct channels and high dimensionality. Meanwhile, the results of this processing could not only substantially increase the computation burden but also result in performance degradation [18,41]. To cope with these issues, a novel event sequence encoding method is proposed as a second symbolization to sufficiently exploit the spatial correlations. This method involves combining and encoding the values from multi-dimensional symbolized sequences at the same moment into a single event symbol. For a D dimension motion sequence, there are at most 3^D possible event cases, which are denoted by $V = \{\beta^1, \beta^2, \dots, \beta^{3^D}\}$. Obviously, the channel number of encoded data is only $1/D$ of the original channel dimension. Figure 2 depicts an example of symbolic representation of 2-dimensional sequence data through the ASC method, where the used data comprise a six-circle gesture motion data from the LIBRAS2 dataset. It can be seen from the figure that the recorded six-circle gesture motion sequence is initially divided into two one-dimensional sequences $X_1^{(n)}$ and $X_2^{(n)}$, and then each sequence is encoded by using the proposed first-order differential symbolization. Finally, all symbolized sequences are converted into a single event sequence as the input of the classification methods.

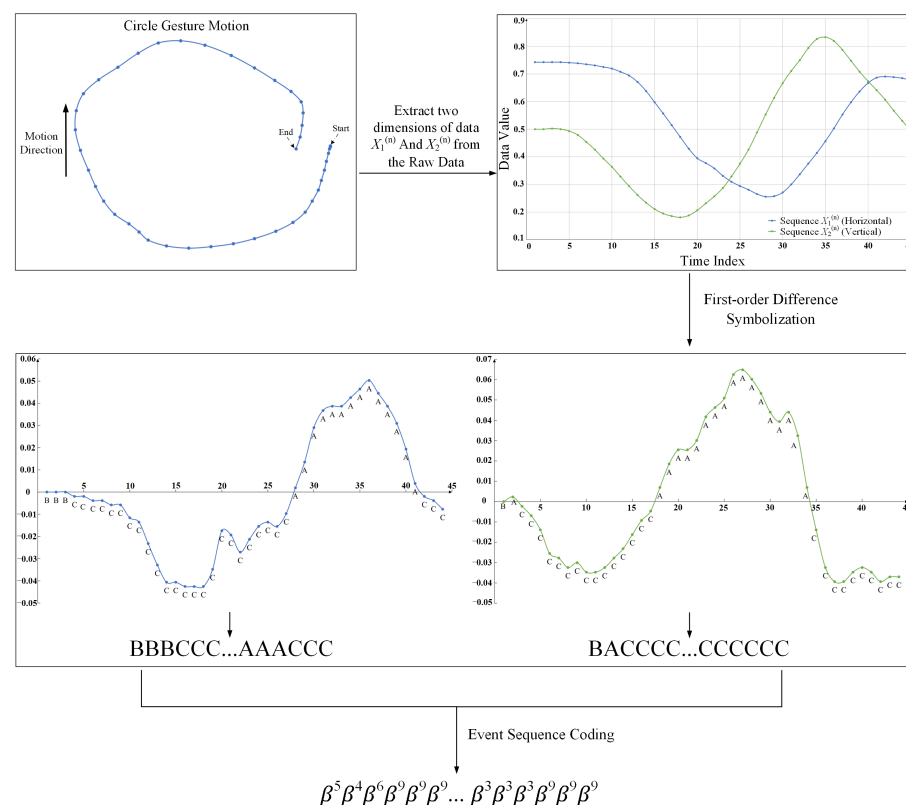


Figure 2. Representation of the six-circle gesture motion sequence in LIBRAS2.

3.3. Ensemble Learning Classification

To build on the above premise, it is necessary to design a model that can effectively process the features finally obtained. Toward this end, a novel HMM-based AdaBoost model is constructed to perform the classification of motion sequences. HMM is a powerful statistical model that can capture the implicit relationships among elements in sequences, namely, the intricate interplay between state transitions and observations within sequence data [42], and provide a comprehensive understanding and representation of chronological feature. At the same time, AdaBoost provides two advantages: firstly, this classifier proficiently classifies multiple classes through the amalgamation of numerous base learners. Secondly, since multi-dimensional motion sequence data usually contain noise and uncertainty; AdaBoost can effectively reduce the risk of overfitting. It will adaptively adjust the weights of the samples and pay more attention to the misclassified samples, thus improving the robustness of the model [26,27].

3.3.1. Constructing Hidden Markov Models

As an extension of Markov chains, HMM is a stochastic process with multiple states. Each state is associated with an observation, and transitions between states are associated with transition probabilities. The key feature of a Markov process is that transitions to the next state depend solely on the current state, independent of the previous other states [24,42]. Since this technology has the property of the unobservability of the real state sequence, it is also referred to as an HMM in the literature.

Let $Q = \{q_1, q_2, \dots, q_Z\}$ be the set of hidden states and M be the number of all event symbols, where Z is the total number of hidden states. An HMM has a parameter set $\lambda = (\Xi, G, \Pi)$ consisting of an initial state matrix Π , a state transition matrix Ξ , and an observation matrix G [42]. The state transition matrix $\Xi(Z \times Z)$ is composed of elements $\xi_{i,j}$, and its specific form is defined as:

$$\Xi = [\xi_{i,j}]_{Z \times Z} \quad (3)$$

$$\xi_{i,j} = P(h_{l+1} = q_j \mid h_l = q_i), 1 \leq i, j \leq Z, l = 1, 2, \dots, L \quad (4)$$

$$\sum_{j=1}^Z \xi_{i,j} = 1 \quad (5)$$

where h_l is a random variable associated with the hidden state of the l -th observed coding symbol in the event sequence, L refers to the length of the event sequence, and $\xi_{i,j}$ represents the probability of transitioning from hidden state q_i to hidden state q_j ($i, j \in \{1, 2, \dots, Z\}$).

Suppose that $g_j(m)$ is the probability of the m -th category event symbol appearing from hidden state q_j ; the observation matrix G can be expressed as:

$$G = [g_j(m)]_{Z \times M} \quad (6)$$

$$g_j(m) = P(o_l = \beta_m \mid h_l = q_j), m = 1, 2, \dots, M \quad (7)$$

$$\sum_{m=1}^M g_j(m) = 1 \quad (8)$$

where o_l stands for a random variable associated with the l -th coding symbol in sequence.

Lastly, the initial state matrix $\Pi(Z \times 1)$ is determined based on the frequency of the initial hidden states in the symbol sequence. Figure 3 presents an illustrative diagram outlining the process for determining the motion category of an event sequence. Specifically,

given any one event sequence, the motion category of the event sequence θ is calculated as follows [23]:

$$P(\theta | \lambda) = \Pi_{h_1} \cdot g_1(o_1) \xi_{1,2}(o_1) g_2(o_2) \cdots \xi_{L-1,L} g_L(o_L) \quad (9)$$

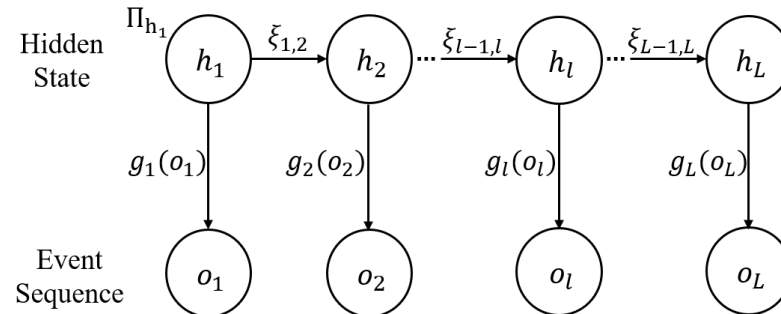


Figure 3. Illustration of the employed event sequence-based HMM.

In this study, each base learner (SequenceHMM) consists of the same number of HMMs as the total number of motion categories in the testing dataset. For instance, the HAR dataset encompasses six distinct motion classes, so the constructed HMM-based learner would contain six HMMs. That is to say, each class constructs a corresponding HMM, denoted by $P(\theta | \lambda_{y=k})$ as the probability of the event sequence θ under the k -th HMM ($k = 1, 2, \dots, K$). Thus, if $P(\theta | \lambda_{y=k})$ exceeds the probability values of the other $K - 1$ HMMs, the event sequence θ_n would be identified as the k -th motion class [23]:

$$\hat{y}(\theta_n) = \arg \max_{k=1,2,\dots,K} P(\theta | \lambda_{y=k}). \quad (10)$$

3.3.2. Constructing Ensemble-SequenceHMM Using AdaBoost

In this step, we deploy an ensemble of SequenceHMM to reflect the diversity of human motion. A single SequenceHMM may effectively classify categories within the training data, while it degenerates in human motion, in which the category is relatively less recorded. Different human motion activities can be influenced by various factors such as environment, age, and scenario. In addition, certain activities with similar movement patterns (e.g., zigzag (vertical) and wavy (vertical)) may be easily confused. Consequently, this work attempts to design an ensemble-SequenceHMM that uses multiple SequenceHMMs with varying hidden states to discriminate between motion sequences, which can enhance model flexibility and effectiveness.

The fundamental principle of the ensemble learning method is that in the learning process, it assigns weights to numerous base learners, which are subsequently amalgamated to form a classifier that can surpass the performance of any individual one [26,27]. The more variable and diverse individual base learners are, the stronger the classification performance of the ensemble learning [23]. Thus, this work constructs several sequence-based HMMs with AdaBoost methodology to classify the categories of the input human motion data.

AdaBoost is an iterative adaptive boosting methodology that can be roughly segmented into three stages, as illustrated in Figure 4. Initially, each sample is assigned equal weight before training the classification model. Then, AdaBoost learns a small number of weak classifiers and iteratively boosts them into a strong classifier with higher precision. In each iteration ϕ ($\phi = 1, 2, \dots, \Phi$), the samples assigned different weights are applied to train the SequenceHMM, and the corresponding base learner weights are updated according to the classification error rate of the event sequences. Subsequently, all sample weights are updated, where the weights of correctly classified samples are decreased and the weights of incorrectly classified samples are increased. Finally, each SequenceHMM with different weights is integrated into the final classifier by a weighting strategy. To sum up, the above process not only pays more attention to the misclassified samples but also inte-

grates multiple SequenceHMMs to capture the correlation information between sequences more effectively, thus improving the overall discriminative ability of the model.

The detailed algorithm for the proposed framework is shown in Algorithm 1. Here, $I(\cdot)$ represents the indicator function. The time complexity of the ASC is $O(3^D L)$. The time complexity of the Adaboost is $O(\Phi K f)$ [43], where f is the complexity of a base learner. In this work, the used base learner is the HMM, and the corresponding time complexity is $O(LZ^2)$ [44]. Therefore, the overall computational complexity of the Algorithm 1 is $O(3^D L + \Phi K L Z^2)$.

Algorithm 1: Pseudo-code for ASC and Ensemble-SequenceHMM

Input: Multi-dimensional motion sequence

$$SDB = \{(X^{(1)}, y_1), (X^{(2)}, y_2), \dots, (X^{(n)}, y_n), \dots, (X^{(N)}, y_N)\}.$$

Initialize: Define number of iterations Φ ;

Initialize all weights over training samples $w_n = \frac{1}{N}$, $n = 1, 2, \dots, N$.

Output: Estimate the class label $\hat{y}$ of each sample:

$$\hat{y} = \underset{k}{\operatorname{argmax}} \{ \sum_{\phi=1}^{\Phi} \alpha^{(\phi)} P(\theta \mid \lambda_{y=1}^{(\phi)}), \dots, \sum_{\phi=1}^{\Phi} \alpha^{(\phi)} P(\theta \mid \lambda_{y=K}^{(\phi)}) \}.$$

```

1 for  $n = 1$  to  $N$  do
2   for  $t = 2$  to  $T$  do
3     Compute first-order differential symbolization by Equations (1) and (2) ;
4   Encode event sequence  $\theta_n$  with  $D$ -dimensional symbol sequences and generate
   sample  $(\theta_n, y_n)$  ;
5  $\Theta \leftarrow \{\theta_1, \theta_2, \dots, \theta_N\}$ ;
6 for  $\phi = 1$  to  $\Phi$  do
7   for  $k = 1$  to  $K$  do
8     Extract  $DB_{y=k}^{(\phi)}$  from  $\Theta$  ;
9     Train  $SequenceHMM^{(\phi)}$ ,  $\lambda_{y=k}^{(\phi)}$ , using  $DB_{y=k}^{(\phi)}$  and weight  $w_n$ , as in
     Section 3.3.1;
10  Estimate the class label  $\hat{y}_n^{(\phi)}$  of each sample by Equation (10);
11  Compute error rate  $err^{(\phi)} = \sum_{n=1}^N w_n \cdot I(\hat{y}^{(\phi)}(\theta_n) \neq y_n)$  ;
12  Compute weight of  $SequenceHMM^{(\phi)}$ ,  $\alpha^{(\phi)} = \log \frac{1-err^{(\phi)}}{err^{(\phi)}} + \log(K-1)$  ;
13  Update weight  $w_n \leftarrow w_n \cdot \exp(\alpha^{(\phi)} \cdot I(\hat{y}^{(\phi)}(\theta_n) \neq y_n))$ ;

```

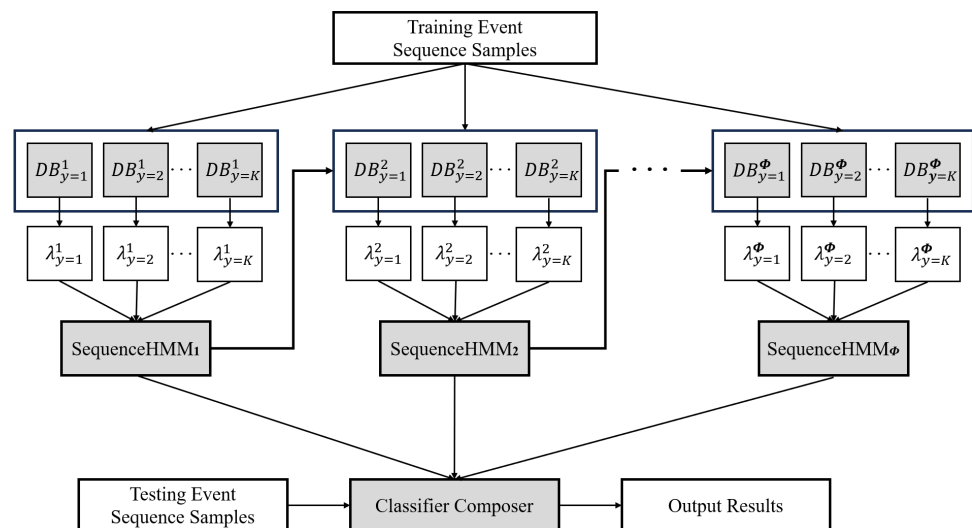


Figure 4. Framework of the proposed Ensemble-SequenceHMM.

4. Experimental Results and Analysis

In this section, we evaluate the effectiveness of the proposed feature representation method on five real-world motion sequence datasets. Meanwhile, we also conduct comparative experiments with several representative approaches. All experiments are conducted on a PC configured with 3.20 GHz, AMD Ryzen 7 5800H CPU, NVIDIA GeForce RTX 3060 GPU, 16 GB RAM, and Windows 11 as the operating system.

4.1. Experimental Setup

To ensure the rigor of our experiment, we selected two mainstream sequence representations for comparison. The first one was SAX [18]. The second one was Adaptive SAX based on Entropy (ASAX_EN) [20], which segments the sequences into different parts according to their entropy and symbolically represents each segment. The processed data of these two feature representations were fed into three different classifiers, namely, KNN, Bayesian, and HMM, to carry out a comprehensive comparison. Note that the input features of KNN and Bayesian are constructed by using six-gram.

Furthermore, we also compared our results against several deep learning models that utilize embedded sequence representations. The first selected model, namely, LSTM, is a variant of a recurrent neural network that can capture intricate dependencies in sequential data. According to the work of Maulik et al. [45] and Tiumentsev et al. [46], the LSTM method generally consists of an LSTM layer with 32 neurons, followed by a fully connected layer. Meanwhile, the Softmax function is used for classification. The second model is the Multilayer Perceptron (MLP), representing an alternative neural network architecture in deep learning. It also has the capability to automatically learn latent features within sequences. Ismail et al. [13] experimented with time series datasets from different domains and showed that a four-layer MLP obtained good overall performance. Due to its success, the MLP used in this work adopted the same network architecture. Additionally, two relatively recent embedded sequence representation methods are also selected as comparative algorithms: Time Le-Net (t-LeNet) [47] and the Neural Network Augmented with Task-Adaptive Projection (TapNet) [12]. The TapNet is a multivariate time series classification method with an attentional prototype network. The t-LeNet is a convolutional neural network based on LeNet, which incorporates specialized structures and layers designed for sequence data to effectively identify and extract features within it.

In all of our experiments, the input length of all deep learning networks is set as the maximum length among all sequences. For our proposed approach, the hidden state number of five SequenceHMMs is set to 4, 5, 6, 7, and 8, respectively. Meanwhile, the Baum–Welch algorithm is applied to learn the model parameters λ of the SequenceHMM from the dataset in this work.

In addition, this work follows a similar experimental protocol to Lima et al. [21] and employs five-fold cross-validation to evaluate the performance of the proposed framework. During five-fold cross-validation, the complete dataset is randomly partitioned into five subsets, where one subset function was used for testing and the rest for training. Subsequently, five experiments were conducted, and the average performance of all iterations was reported.

4.2. Evaluation Metrics

To evaluate the proposed framework, we adopt two widely used classification metrics. The first metric is overall accuracy, which is capable of measuring the classification accuracy of the classifier over the entire multi-class dataset. The second metric is Macro-F1 [25], which integrates precision and recall for each motion category, providing a comprehensive evaluation of the model's performance through the computation of their harmonized

average. For each motion class, F1-measure is used for evaluation. The calculation of these metrics can be expressed using the following mathematical formulations:

$$\text{Overall accuracy} = \frac{\sum (1|y_n = Py_n)}{\sum (1|y_n = y_n)} \quad (11)$$

$$F1 = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (12)$$

$$\text{Macro-F1} = \frac{\sum_{j=1}^K F1_j}{K} \quad (13)$$

where y_n and Py_n stand for the true and predicted label of the n -th sample, respectively. *Recall* is the ratio of accurate positive predictions to the total number of positive samples. *Precision* is the proportion of real positive samples recognized as positive samples. $F1_j$ is the F1 score of j -th class.

4.3. Overall Comparison with Previous Studies

To illustrate the performance of the proposed framework, we conducted a comparative analysis using five real-world motion sequence datasets. Tables 2 and 3 summarize the comparison between our framework and other methodologies on different datasets. We can discover that our framework outperforms the two mainstream sequence representation approaches. In terms of accuracy, the proposed framework surpasses SAX coupled with N-gram+Bayes by 34.17%, 30.00%, 56.89%, 27.58%, and 55.92% on the LIBRAS1, LIBRAS2, HAR, JSI, and OPPORTUNITY datasets, respectively. In terms of Macro-F1 measure, the proposed framework improved the baseline (SAX + HMM) by more than 35% in all datasets used. Similar results can also be found in comparison with other SAX-based methods. Through in-depth analysis, we think the performance improvement comes mainly from the following two key advantages. On the one hand, the first-order difference symbolization in the proposed ASC can effectively perform both denoising and data smoothing through the differencing of adjacent data points; its ability to learn correlations between sequence elements is an essential aspect that is especially overlooked by other existing representation approaches. On the other hand, adaptive event sequence coding can significantly compress data dimensionality and facilitate the more efficient capture of structural relationships among motion sequences, thus enhancing subsequent classification performance. Furthermore, it is well known that some existing conventional classifiers (e.g., Naive Bayes) often need to rely on the independence assumption—that elements in sequences are independent of each other, to ensure mining efficiency—but this assumption might be violated in practical scenarios [48]. Conversely, the base learner HMM used can capture the dependencies among elements in sequences [23]. Interestingly, it can also be observed from the Tables 2 and 3 that the HMM-based methods generally achieved better classification performance compared to the Bayes-based methods.

Also, when compared against deep learning methods, the proposed framework matches the highest overall accuracy and Macro-F1 on the LIBRAS1, LIBRAS2, and HAR datasets, respectively. Particularly, the overall accuracy of our method on the LIBRAS1, LIBRAS2, HAR, JSI, and OPPORTUNITY datasets demonstrates notable improvements of 14.35%, 8.63%, 67.42%, 40.13%, and 38.00%, respectively, as opposed to the TapNet model. Likewise, the proposed approach outperforms t-LeNet in both evaluation metrics, where the latter obtained suboptimal accuracies on four datasets, except for the LIBRAS2 dataset. This discrepancy can be attributed to the deep learning method's demand for substantial data for effective model training, and it often falls short of achieving satisfactory results when dealing with small sample datasets. The complex model with large-scale parameters is another key for the performance of the deep learning models [13]. Conversely, our proposed framework can achieve satisfactory results without such requirements. Intriguingly, this finding aligns with the results reported in the literature [49,50]. Furthermore,

our proposal enhances data interpretability due to the fact that a finite set of symbols is adapted to represent original numerical values, making the data more understandable and interpretable for further analysis. Consequently, in the context of analyzing small-size motion datasets with small sample, this work focuses on machine learning techniques instead of deep learning.

Table 2. Comparison of overall accuracy (%) with different methods using multiple same public datasets.

Feature Representation	Classifier	LIBRAS1	LIBRAS2	HAR	JSI	OPPORTUNITY
SAX	N-gram+KNN	56.67	51.67	22.00	43.08	27.52
	N-gram+Bayes	60.00	58.33	30.75	37.05	20.22
	HMM	54.17	43.33	31.50	46.84	38.16
ASAX_EN	N-gram+KNN	63.33	54.17	33.25	43.85	50.60
	N-gram+Bayes	66.67	66.67	23.25	34.66	51.26
	HMM	66.67	57.50	39.25	49.67	51.26
Embedded	LSTM	20.25	21.01	21.25	40.98	37.86
	MLP	25.29	16.85	29.75	66.48	74.93
	t-LeNet	93.33	76.52	58.25	42.15	52.45
	TapNet	79.82	78.95	24.50	38.86	38.14
ASC	AdaBoost	94.17	88.33	87.64	64.63	76.14

The best results are highlighted in bold.

Table 3. Comparison of Macro-F1 (%) with different methods using multiple same public datasets.

Feature Representation	Classifier	LIBRAS1	LIBRAS2	HAR	JSI	OPPORTUNITY
SAX	N-gram+KNN	54.61	44.97	13.93	27.98	10.61
	N-gram+Bayes	59.03	54.61	18.28	20.56	8.41
	HMM	52.06	37.60	18.73	24.81	13.81
ASAX_EN	N-gram+KNN	61.62	51.05	31.95	39.34	34.71
	N-gram+Bayes	66.52	64.02	17.78	33.26	34.89
	HMM	65.25	54.06	34.82	40.14	34.89
Embedded	LSTM	8.17	14.12	9.98	14.51	13.74
	MLP	14.73	10.78	20.75	54.05	70.10
	t-LeNet	93.28	73.19	56.24	16.31	39.44
	TapNet	81.55	78.69	20.14	34.08	32.71
ASC	AdaBoost	93.99	87.32	87.56	60.68	61.64

The best results are highlighted in bold.

In conclusion, these results validate the effectiveness of our proposal and highlight the potential advantages of learning inter-correlations among elements and sequences and implementing efficient dimensionality reduction, thereby improving subsequent classification performance. In conjunction with the analysis of the experimental results, our proposed machine learning framework is more tailored for handling multi-dimensional small sample sequence data. This is in contrast to deep learning methods, which demand substantial amounts of training data to ensure their discriminative capacity. Thus, this work chooses to construct a new machine learning framework to address the challenges of motion sequence analysis.

4.4. Ablation Experiments

In this subsection, we provide two additional ablation experiments in five real-life human motion datasets, aiming to conduct an in-depth analysis. The primary experiment is to validate the capability of the presented feature representation through the evaluation of the dual symbolization strategy. The secondary experiment seeks to authenticate the

dependability of ensemble learning. This is achieved by comparing the proposed AdaBoost classifier with a single SequenceHMM based on different hidden states.

4.4.1. Impact of Adaptive Motion Sequence Coding

ASC employs a dual symbolization strategy for fully exploring potential intrinsic structural relationships between sequences. To comprehensively assess the performance of the ASC on motion sequence detection, we have chosen the ASC method without the event sequence coding, and we have used the SAX to replace first-order differential symbolization, which are named model A and model B, respectively. Here, we refer to the first step of symbolization as Symbolization I. Table 4 shows the experimental results obtained on various datasets.

Table 4. Ablation study classification accuracy (%) of the proposed ASC method.

Feature Representation	Symbolization I	Event Sequence Coding	LIBRAS1	LIBRAS2	HAR	JSI	OPPORTUNITY
Model A	✓		59.17	51.67	57.85	49.89	50.03
Model B	✓ (SAX)	✓	88.33	75.83	52.50	59.81	70.16
ASC	✓	✓	94.17	88.33	87.64	64.63	76.14

The best results are highlighted in bold.

As shown in Table 4, when compared to the ASC, the accuracies of Model A on the LIBRAS1, LIBRAS2, and HAR datasets significantly decrease from 94.17%, 88.33%, and 87.64% to 59.17%, 51.67%, and 57.85%, respectively, which indicates the importance of hidden structural information among multi-dimensional sequences to improve classification performance. These results support that the designed event sequence encoding can be a reasonable strategy to identify the motion pattern with efficiency and reliability. In addition, when compared to Model B, the accuracy of the ASC has 5.84%, 12.50%, 35.14%, 4.82%, and 5.98% improvements on the LIBRAS1, LIBRAS2, HAR, JSI, and OPPORTUNITY datasets, respectively. The comparison means that the proposed first-order difference symbolization can characterize sequence data more efficiently and optimize subsequent analysis performance. This can be attributed to the fact that our method utilizes the first-order difference to filter out random fluctuations and discern the correlation between elements (i.e., variant trends), which is key for characterizing motion sequences. More importantly, it also effectively mitigates the effects of noise or outliers in the sequences. In a nutshell, the above result showcases that the proposed dual symbolization mechanism may not only be highly optimal but also plays an important role in constructing a more compact representation, which enables the classification method to gain better performance with limited training data.

4.4.2. Impact of Ensemble Learning

The classification of multi-dimensional sequence data often involves various types of motion activities and exhibits an imbalance in the dataset across different data classes. Therefore, the use of an AdaBoost classifier, which combines multiple base learners, can significantly enhance the classification performance. In order to further investigate the effect of a single base learner versus an ensemble learning classifier on the classification performance of motion sequences, five single SequenceHMMs as classifiers, whose hidden states were set as 4, 5, 6, 7, and 8, respectively, were applied to test and compare with the AdaBoost classifier proposed in this work. This ablation experiment assessed overall accuracy across the five datasets mentioned before.

As shown in Figure 5, the ensemble learning method achieved the best performance in classifying motion sequences, where the HSN denotes the hidden state number of the corresponding SequenceHMM. In comparison to the suboptimal results (HSN = 8), our method exhibits noteworthy enhancements in overall accuracy, with improvements of 4.17%, 0.83%, 1.25%, 11.45%, and 9.57% on the LIBRAS1, LIBRAS2, HAR, JSI, and OP-

PORTUNITY datasets, respectively, which demonstrates the effectiveness of constructing ensemble-sequenceHMM using AdaBoost for motion sequence detection. This is one of the main reasons why we adopt the ensemble learning technique to solve the multiple classification problem of motion sequences in this paper. To our knowledge, it can mitigate the limitations of individual models by combining predictions from multiple base classifiers, resulting in a more robust and accurate prediction. Specifically, it can adapt to data diversity, mitigate the risk of overfitting, and improve the generalization capabilities on new data, based on a collective decision. Moreover, ensemble learning could typically show better performance compared to other machine learning approaches when dealing with unbalanced data [51]. Therefore, this work will provide valuable insights for research aiming to enhance the performance of motion sequence classification systems.

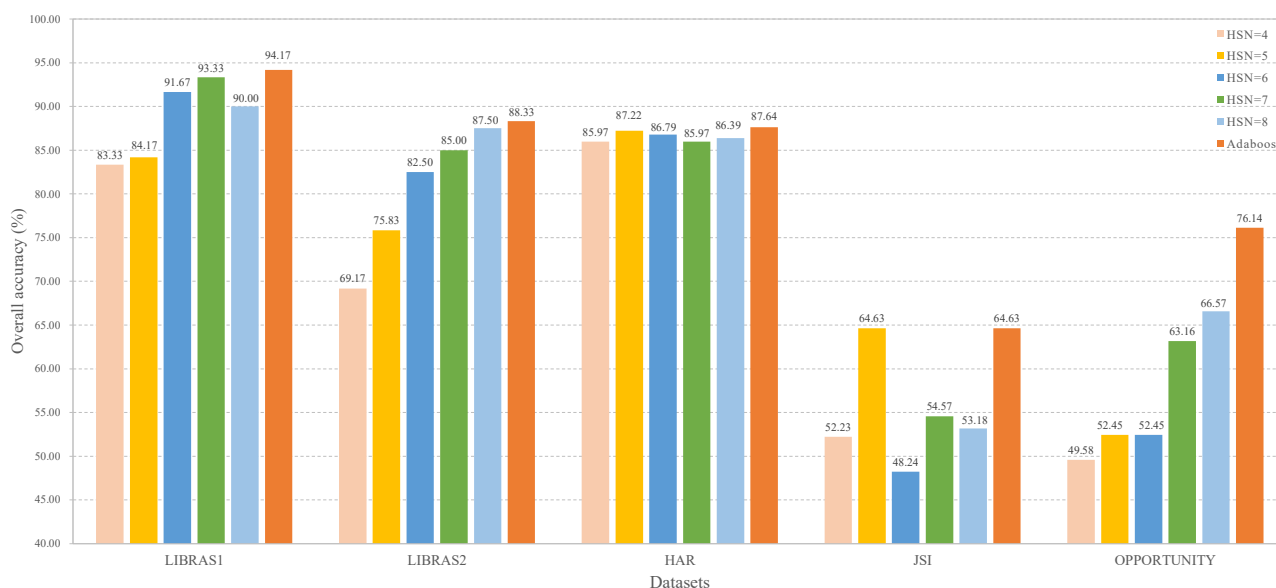


Figure 5. Ablation study classification accuracy (%) of the proposed AdaBoost method.

5. Conclusions and Future Scope

In this study, we propose an adaptive coding feature representation methodology for multi-dimensional motion sequence analysis, which is a critical yet challenging task in motion recognition field. The novel feature representation model, namely, ASC, can sufficiently achieve efficient dimensionality reduction and noise reduction with the help of first-order difference and dual symbolization. To the best of our knowledge, no similar methodologies have been applied to motion sequence detection in the literature. The experiments on real-life human motion datasets show that our proposed framework outperforms rival methodologies, indicating the potential effectiveness of our approach. Additionally, the ablation experiments provide supporting evidence for the superiority of the dual symbolization mechanism and ensemble learning in enhancing the performance of motion data.

Despite the strength of the proposed framework, the discussion is incomplete without mentioning its limitations. Through in-depth analysis of our method, it can be observed that an exponential relationship exists between the sequence dimension and the size of possible event cases. Excessively high dimensions of sequences may result in reduced coding efficiency and increased computational cost. One possible future development of this work involves mitigating the risk of rapid expansion of event cases, via incorporating an unsupervised adaptive grouping strategy. It will further partition highly similar event cases into the same group and encode distinct groups with a finite number of new cases. Another aspect of future work can be enhancing first-order difference symbolization by optimizing spatially sensitive values using local feature statistics or other data distribution-based statistical analysis strategies.

Author Contributions: Conceptualization, X.K., W.K. and T.W.; methodology, W.K., X.K. and X.L.; software, S.C. and T.W.; validation, W.K., Z.L., J.C. and T.W.; formal analysis, X.K. and T.W.; investigation, S.C. and T.W.; resources, W.K. and Z.L.; data curation, X.K. and X.L.; writing—original draft preparation, X.K. and X.L.; writing—review and editing, Z.L., W.K., J.C., X.K. and X.L.; visualization, S.C., Z.L. and T.W.; supervision, Z.L. and T.W.; project administration, X.K.; and funding acquisition, X.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key Research and Development Program of China (No. 2020YFF0401865, No. 2021YFF1200700, No. 2023YFF1203900) and the Marine Aquaculture and Intelligent IOT Technology Innovation Research Team Funding, Fujian Agriculture and Forestry University.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: <https://archive.ics.uci.edu/>.

Acknowledgments: The authors would like to thank all the anonymous reviewers for their insightful comments and constructive suggestions that have obviously upgraded the quality of this manuscript.

Conflicts of Interest: The authors declare that they have no known competing financial interest or personal circumstances that could have appeared to influence the work reported in this manuscript.

References

1. Yao, C.; He, J.S.; Che, H.; Huang, Y.; Wu, J. Feature pyramid self-attention network for respiratory motion prediction in ultrasound image guided surgery. *Int. J. Comput. Assist. Radiol. Surg.* **2022**, *17*, 2349–2356. [\[CrossRef\]](#) [\[CrossRef\]](#) [\[PubMed\]](#)
2. Franco, T.; Sestrem, L.; Henriques, P.R.; Alves, P.; Varanda Pereira, M.J.; Brandão, D.; Leitão, P.; Silva, A. Motion sensors for knee angle recognition in muscle rehabilitation solutions. *Sensors* **2022**, *22*, 7605. [\[CrossRef\]](#) [\[CrossRef\]](#) [\[PubMed\]](#)
3. Chavarriaga, R.; Sagha, H.; Calatroni, A.; Digumarti, S.T.; Tröster, G.; Millán, J.; Roggen, D. The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognit. Lett.* **2013**, *34*, 2033–2042. [\[CrossRef\]](#)
4. Kulakou, S.; Ragab, N.; Midoglu, C.; Boeker, M.; Johansen, D.; Riegler, M.A.; Halvorsen, P. Exploration of Different Time Series Models for Soccer Athlete Performance Prediction. *Eng. Proc.* **2022**, *18*, 37. [\[CrossRef\]](#)
5. Hao, W.W. Classification of Sport Actions Using Principal Component Analysis and Random Forest Based on Three-Dimensional Data. *Displays* **2022**, *72*, 102135. [\[CrossRef\]](#)
6. Nguyen, T.; Nguyen, G.; Nguyen, B.M. EO-CNN: An Enhanced CNN Model Trained by Equilibrium Optimization for Traffic Transportation Prediction. *Procedia Comput. Sci.* **2020**, *176*, 800–809. [\[CrossRef\]](#) [\[CrossRef\]](#)
7. Majumder, A.J.; Izaguirre, J.A. A Smart IoT Security System for Smart-Home Using Motion Detection and Facial Recognition. In Proceedings of the 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), Madrid, Spain, 13–17 July 2020; pp. 1065–1071. [\[CrossRef\]](#)
8. Chen, J.F.; Wang, C.C.; Wu, E.H.K.; Chou, C.F. Simultaneous heterogeneous sensor localization, joint tracking, and upper extremity modeling for stroke rehabilitation. *IEEE Syst. J.* **2020**, *14*, 3570–3581. [\[CrossRef\]](#) [\[CrossRef\]](#)
9. Qiu, Y.Y.; Guan, Y.R.; Liu, S. The analysis of infrared high-speed motion capture system on motion aesthetics of aerobics athletes under biomechanics analysis. *PLoS ONE* **2023**, *18*, e0286313. [\[CrossRef\]](#) [\[CrossRef\]](#)
10. Chen, L.F.; Wu, H.Y.; Kang, W.X.; Wang, S.R. Symbolic sequence representation with Markovian state optimization. *Pattern Recognit.* **2022**, *131*, 108849. [\[CrossRef\]](#) [\[CrossRef\]](#)
11. Huang, B.Z.; Li, X.D. Human Motion Prediction via Dual-Attention and Multi-Granularity Temporal Convolutional Networks. *Sensors* **2023**, *23*, 5653. [\[CrossRef\]](#) [\[CrossRef\]](#)
12. Yoon, S.W.; Seo, J.; Moon, J. TapNet: Neural Network Augmented with Task-Adaptive Projection for Few-Shot Learning. In Proceedings of the 36th International Conference on Machine Learning (ICML 2019), Long Beach, CA, USA, 9–15 June 2019; pp. 7115–7123. [\[CrossRef\]](#)
13. Ismail Fawaz, H.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.A. Deep learning for time series classification: A review. *Data Min. Knowl. Discov.* **2019**, *33*, 917–963. [\[CrossRef\]](#) [\[CrossRef\]](#)
14. Bai, B.; Li, G.L.; Wang, S.Z.; Wu, Z.D.; Yan, W.H. Time series classification based on multi-feature dictionary representation and ensemble learning. *Expert Syst. Appl.* **2021**, *169*, 114162. [\[CrossRef\]](#) [\[CrossRef\]](#)
15. Buchaiah, S.; Shakya, P. Bearing fault diagnosis and prognosis using data fusion based feature extraction and feature selection. *Measurement* **2022**, *188*, 110506. [\[CrossRef\]](#) [\[CrossRef\]](#)
16. Katarya, R.; Prasad, T. A Survey on Time Series Online Sequential Learning Algorithms. In Proceedings of the 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICIC), Coimbatore, India, 14–16 December 2017; pp. 1–4. [\[CrossRef\]](#)
17. Chen, R.J.; Ravishanker, N. Feature Construction Using Persistence Landscapes for Clustering Noisy IoT Time Series. *Future Internet* **2023**, *15*, 195. [\[CrossRef\]](#) [\[CrossRef\]](#)

18. Lin, J.; Keogh, E.; Wei, L.; Lonardi, S. Experiencing SAX: A novel symbolic representation of time series. *Data Min. Knowl. Discov.* **2007**, *15*, 107–144. [\[CrossRef\]](#) [\[CrossRef\]](#)
19. Li, Y.C.; Shen, D.R. A new symbolic representation method for time series. *Inf. Sci.* **2022**, *609*, 276–303. [\[CrossRef\]](#) [\[CrossRef\]](#)
20. Djebour, L.; Akbarinia, R.; Massegia, F. Variable-Size Segmentation for Time Series Representation. In *Transactions on Large-Scale Data- and Knowledge-Centered Systems LIII*; Abdelkader, H., A Min, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2023; pp. 34–65. [\[CrossRef\]](#)
21. Sousa Lima, W.; de Souza Bragança, H.L.; Montero Quispe, K.G.; Pereira Souto, E.J. Human activity recognition based on symbolic representation algorithms for inertial sensors. *Sensors* **2018**, *18*, 4045. [\[CrossRef\]](#) [\[CrossRef\]](#)
22. Le, Nguyen, T.; Gsponer, S.; Ilie, I.; O'reilly, M.; Ifrim, G. Interpretable time series classification using linear models and multi-resolution multi-domain symbolic representations. *Data Min. Knowl. Discov.* **2019**, *33*, 1183–1222. [\[CrossRef\]](#) [\[CrossRef\]](#)
23. Kang, M.; Ahn, J.; Lee, K. Opinion mining using ensemble text hidden Markov models for text classification. *Expert Syst. Appl.* **2018**, *94*, 218–227. [\[CrossRef\]](#) [\[CrossRef\]](#)
24. Gomes, S.R.; Saroar, S.G.; Mosfaiul, M.; Telot, A.; Khan, B.N.; Chakrabarty, A.; Mostakim, M. A comparative approach to email classification using Naive Bayes classifier and hidden Markov model. In Proceedings of the 2017 4th International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, Bangladesh, 28–30 September 2017; pp. 482–487. [\[CrossRef\]](#)
25. Dogan, T.; Uysal, A.K. A novel term weighting scheme for text classification: TF-MONO. *J. Inf.* **2020**, *14*, 101076. [\[CrossRef\]](#) [\[CrossRef\]](#)
26. Hastie, T.; Rosset, S.; Zhu, J.; Zou, H. Multi-class adaboost. *Stat. Its Interface* **2009**, *2*, 349–360. [\[CrossRef\]](#) [\[CrossRef\]](#)
27. Ding, J.Y.; Wang, Y.; Si, H.Y.; Gao, S.; Xing, J.W. Multimodal Fusion-AdaBoost Based Activity Recognition for Smart Home on WiFi Platform. *IEEE Sensors J.* **2022**, *22*, 4661–4674. [\[CrossRef\]](#) [\[CrossRef\]](#)
28. Agrawal, R.; Faloutsos, C.; Swami, A. Efficient similarity search in sequence databases. In *Foundations of Data Organization and Algorithms, Proceedings of the 4th International Conference, FODO '93, Chicago, IL, USA, 13–15 October 1993*; Proceedings; Springer: Berlin/Heidelberg, Germany, 1993; pp. 69–84. [\[CrossRef\]](#)
29. Yang, W.K.; Du, Q.L.; Cui, J.C.; Wang, Y.K.; Lu, X.P.; Qi, C.X.; Zhang, G.P. Motion recognition based on sum of the squared errors distribution. *IEEE Access* **2021**, *9*, 37116–37130. [\[CrossRef\]](#) [\[CrossRef\]](#)
30. Liu, X.; Liu, H.; Guo, Q.; Zhang, C. Adaptive wavelet transform model for time series data prediction. *Soft Comput.* **2020**, *24*, 5877–5884. [\[CrossRef\]](#) [\[CrossRef\]](#)
31. Chatzigeorgakidis, G.; Skoutas, D.; Patroumpas, K.; Palpanas, T.; Athanasiou, S.; Skiadopoulos, S. Efficient range and knn twin subsequence search in time series. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 5794–5807. [\[CrossRef\]](#) [\[CrossRef\]](#)
32. Zhang, X.; Sun, Y.C.; Zhu, H.Q. Method of Pilot's Motion Retrieval and Recognition Based on Dynamic Programming. In Proceedings of the 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), Shanghai, China, 16–18 August 2017; pp. 345–349. [\[CrossRef\]](#)
33. Keogh, E.J.; Pazzani, M.J. An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback. In Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), New York, NY, USA, 27–31 August 1998; pp. 239–243. [\[CrossRef\]](#)
34. Kang, W.X.; Chen, L.F.; Guo, G.D. Spatio-temporal structure feature representation model for motion sequences. *CAAI Trans. Intell. Syst.* **2023**, *18*, 240–250. [\[CrossRef\]](#)
35. Junejo, I.N.; Junejo, K.N.; Aghbari, Z.A. Silhouette-based human action recognition using SAX-Shapes. *Vis. Comput.* **2014**, *30*, 259–269. [\[CrossRef\]](#) [\[CrossRef\]](#)
36. Zhou, T.Y.; Wang, Y.B.; Du, J. Human Intent Prediction in Human-Robot Collaboration—A Pipe Maintenance Example. In Proceedings of the Construction Research Congress 2022, Arlington, VA, USA, 9–12 March 2022; pp. 581–590. [\[CrossRef\]](#)
37. Riza, L.S.; Fazanadi, M.N.; Utama, J.A.; Samah, K.A.F.A.; Hidayat, T.; Nazir, S. SAX and Random Projection Algorithms for the Motif Discovery of Orbital Asteroid Resonance Using Big Data Platforms. *Sensors* **2022**, *22*, 5071. [\[CrossRef\]](#) [\[CrossRef\]](#)
38. Dias, D.B.; Madeo, R.C.B.; Rocha, T.; Biscaro, H.H.; Peres, S.M. Hand movement recognition for Brazilian Sign Language: A study using distance-based neural networks. In Proceedings of the 2009 International Joint Conference on Neural Networks (IJCNN), Atlanta, GA, USA, 14–19 June 2009; pp. 697–704. [\[CrossRef\]](#)
39. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. A public domain dataset for human activity recognition using smartphones. In Proceedings of the 21st European Symposium on Artificial Neural Networks (ESANN 2013), Bruges, Belgium, 24–26 April 2013; p. 3. [\[CrossRef\]](#)
40. Sprint, G.; Cook, D.; Weeks, D.; Dahmen, J.; La Fleur, A. Analyzing Sensor-Based Time Series Data to Track Changes in Physical Activity during Inpatient Rehabilitation. *Sensors* **2017**, *17*, 2219. [\[CrossRef\]](#) [\[CrossRef\]](#)
41. Gu, J.X.; Yu, Z.W.; Shen, K.L. Alohoma: Motion-Based Hotword Detection in Head-Mounted Displays. *IEEE Internet Things J.* **2020**, *7*, 611–620. [\[CrossRef\]](#) [\[CrossRef\]](#)
42. Ishibashi, N.; Fujii, F. Hidden Markov model-based human action and load classification with three-dimensional accelerometer measurements. *IEEE Sensors J.* **2020**, *21*, 6610–6622. [\[CrossRef\]](#) [\[CrossRef\]](#)
43. Sun, J.W.; Zou, R.; Liang, R.X.; Gao, L.; Liu, S.; Li, Q.; Zhang, K.; Jiang, L.L. Ensemble Knowledge Tracing: Modeling interactions in learning process. *Expert Syst. Appl.* **2022**, *207*, 117680. [\[CrossRef\]](#) [\[CrossRef\]](#)
44. Castellini, A.; Masillo, F.; Azzalini, D.; Amigoni, F.; Farinelli, A. Adversarial Data Augmentation for HMM-based Anomaly Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 14131–14143. [\[CrossRef\]](#) [\[CrossRef\]](#) [\[PubMed\]](#)

45. Maulik, R.; Mohan, A.; Lusch, B.; Madireddy, S.; Balaprakash P.; Livescu D. Time-series learning of latent-space dynamics for reduced-order model closure. *Phys. D Nonlinear Phenom.* **2020**, *405*, 132368. [[CrossRef](#)] [[CrossRef](#)]
46. Tiumentsev, A.Y.; Tiumentsev, Y.V. Motion Control of Supersonic Passenger Aircraft Using Machine Learning Methods. *Opt. Mem. Neural Networks* **2023**, *32*, S195–S205. [[CrossRef](#)] [[CrossRef](#)]
47. Le Guennec, A.; Malinowski, S.; Tavenard, R. Data augmentation for time series classification using convolutional neural networks. In Proceedings of the ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, Riva Del Garda, Italy, 19–23 September 2016. [[CrossRef](#)]
48. Xing, Z.Z.; Pei, J.; Keogh, E. A brief survey on sequence classification. *ACM SIGKDD Explor. Newsl.* **2010**, *12*, 40–48. [[CrossRef](#)] [[CrossRef](#)]
49. Su, H.; Lu, X.M.; Chen, Z.Q.; Zhang, H.S.; Lu, W.F.; Wu, W.T. Estimating Coastal Chlorophyll-A Concentration from Time-Series OLCI Data Based on Machine Learning. *Remote Sens.* **2021**, *13*, 576. [[CrossRef](#)] [[CrossRef](#)]
50. Zhang, Y.; Ren, G.; Liu, X.J.; Gao, G.L.; Zhu, M.D. Ensemble learning-based modeling and short-term forecasting algorithm for time series with small sample. *Eng. Rep.* **2022**, *4*, e12486. [[CrossRef](#)] [[CrossRef](#)]
51. Jain, R.; Ganesan, R.A. Reliable sleep staging of unseen subjects with fusion of multiple EEG features and RUSBoost. *Biomed. Signal Process. Control* **2021**, *70*, 103061. [[CrossRef](#)] [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.