



Article

Automated Student Classroom Behaviors' Perception and Identification Using Motion Sensors

Hongmin Wang ^{1,†}, Chi Gao ^{2,3,†}, Hong Fu ^{1,*}, Christina Zong-Hao Ma ^{4,*}, Quan Wang ², Ziyu He ¹ and Maojun Li ^{1,5}

¹ Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong 999077, China

² The Key Laboratory of Spectral Imaging Technology, Xi'an Institute of Optics and Precision Mechanics of the Chinese Academy of Sciences, Xi'an 710119, China

³ The University of Chinese Academy of Sciences, Beijing 100049, China

⁴ Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong 999077, China

⁵ School of Information Science and Technology, Northwest University, Xi'an 710127, China

* Correspondence: hfu@eduhk.hk (H.F.); czh.ma@polyu.edu.hk (C.Z.-H.M.); Tel.: +852-2948-7535

† These authors contributed equally to this work.

Abstract: With the rapid development of artificial intelligence technology, the exploration and application in the field of intelligent education has become a research hotspot of increasing concern. In the actual classroom scenarios, students' classroom behavior is an important factor that directly affects their learning performance. Specifically, students with poor self-management abilities, particularly specific developmental disorders, may face educational and academic difficulties owing to physical or psychological factors. Therefore, the intelligent perception and identification of school-aged children's classroom behaviors are extremely valuable and significant. The traditional method for identifying students' classroom behavior relies on statistical surveys conducted by teachers, which incurs problems such as being time-consuming, labor-intensive, privacy-violating, and an inaccurate manual intervention. To address the above-mentioned issues, we constructed a motion sensor-based intelligent system to realize the perception and identification of classroom behavior in the current study. For the acquired sensor signal, we proposed a Voting-Based Dynamic Time Warping algorithm (VB-DTW) in which a voting mechanism is used to compare the similarities between adjacent clips and extract valid action segments. Subsequent experiments have verified that effective signal segments can help improve the accuracy of behavior identification. Furthermore, upon combining with the classroom motion data acquisition system, through the powerful feature extraction ability of the deep learning algorithms, the effectiveness and feasibility are verified from the perspectives of the dimensional signal characteristics and time series separately so as to realize the accurate, non-invasive and intelligent children's behavior detection. To verify the feasibility of the proposed method, a self-constructed dataset (SCB-13) was collected. Thirteen participants were invited to perform 14 common class behaviors, wearing motion sensors whose data were recorded by a program. In SCB-13, the proposed method achieved 100% identification accuracy. Based on the proposed algorithms, it is possible to provide immediate feedback on students' classroom performance and help them improve their learning performance while providing an essential reference basis and data support for constructing an intelligent digital education platform.



Citation: Wang, H.; Gao, C.; Fu, H.; Ma, C.Z.-H.; Wang, Q.; He, Z.; Li, M. Automated Student Classroom Behaviors' Perception and Identification Using Motion Sensors. *Bioengineering* **2023**, *10*, 127. <https://doi.org/10.3390/bioengineering10020127>

Academic Editor: Franz Konstantin Fuss

Received: 26 December 2022

Revised: 11 January 2023

Accepted: 12 January 2023

Published: 18 January 2023



Copyright: © 2023 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/>).

Keywords: intelligent system; deep learning; classroom behavior; motion identification

1. Introduction

1.1. Background Information on Students' Classroom Behavior

With the rapid development, penetration, and integration of artificial intelligence technologies in various areas of society, intelligent digital-based education is progressively becoming a hot issue of substantive research [1,2]. Among the many educational research

carriers, the intelligent education classroom scenarios are still the commonly adopted educational method [3], which has the outstanding advantages of direct feedback and extensive interaction between teachers and students [4].

Classroom scenarios present complexity and diversity according to different participants and instructional content. Research has shown that in classroom scenarios, students' classroom behavior is one of the most important factors influencing their academic performance [5]. Compared with high-achieving students, low-achieving students typically spend a significant amount of class time engaged in non-academic work or other academic work [6]. Therefore, investigating student classroom behavior has essential research implications and applicate values for enhancing student performance and promoting instructional strategies [7].

Specifically, the study of classroom behavior through detecting and identifying students' classroom behavior patterns can provide timely and stage-specific feedback on students' classroom performance. Effective statistical analysis of students' behavior patterns will assist students in effectively understanding their learning habits, timely correcting their poor classroom behavior, improving learning strategies, adjusting learning progress, and deepening their understanding and absorption of knowledge.

Furthermore, the analysis of students' classroom behavior is especially beneficial for students with special education needs (SEN) and developmental disabilities, such as attention deficit and hyperactivity disorder (ADHD) [8], autism spectrum disorder (ASD) [9], and learning disabilities [9,10]. Conducting classroom behavior analysis is crucial to improving these students' classroom performance and enhancing their classroom concentration. The percentage of school-aged children diagnosed with developmental disorders is increasing dramatically each year due to various environmental factors such as location, level of education, and medical care. In addition, the percentage of children with developmental disorders increased to 17.8% of all children (3–17 years old, the United States). The proportion is substantial, with approximately one in six children diagnosed with a disease [11]. Specifically, ADHD has the broadest range of effects on all developmental disorders and has the most significant prevalence among children. Characteristics of children with ADHD include inattention, hyperactivity, and impulsivity. Students with developmental disorders generally suffer from academic problems due to physical or psychological issues, and their classroom performance is difficult to self-control.

The study of classroom behaviors of students with developmental disabilities can be used to detect and identify their classroom behaviors automatically to a large extent. It can help them improve their self-awareness, enhance their concentration, and effectively achieve supplementary education without external interventions [12]. Auxiliary education based on non-artificial reminders can greatly relieve their learning pressure, ease learning difficulties and anxiety, increase knowledge and improve environmental adaptability, and promote a virtuous cycle of learning [13].

Finally, due to the need to build intelligent digital education platforms for schools and parents, the study of classroom behavior can further refine students' learning performance at school [14], optimize school teaching services, improve teaching strategies, and facilitate communication and exchange among multiple parties [15]. The intelligent digital platform is designed with students as the primary body and their classroom behaviors as the principal way of measuring their classroom status in order to enhance students' learning performance and optimize the teaching services of teachers. The perception and identification of students' classroom behaviors open the door to the development of an intelligent digital education platform.

1.2. Literature Review

In this part, we review the literature from the perspectives of 'Existing methods and the limitations' and 'Advanced methods on human activity recognition' to demonstrate the previous work on the perception and identification of student classroom behaviors.

1.2.1. Existing Methods and the Limitations

The previous research on students' classroom behavior in the traditional education field is often based on statistical survey methods, requiring teachers to observe the classroom behavior of the entire class or a smaller number of people as an observer over a period within the classroom and to record their behavior [16]. In these circumstances, the teacher plays the role of the evaluator to assess the student's behavior patterns. This manual, vision-based approach is usually surpassed in identifying inappropriate classroom activities, but the teacher's one-to-many nature at the time of the count results in the poor perception of the finer classroom behaviors of most students [17]. Furthermore, this visual-based artificial approach to behavior analysis is undoubtedly time-consuming, labor-intensive, and highly subjective. It is highly likely to violate students' privacy through external interventions and create learning ancients. It cannot give objective science-based judgments, thus preventing a comprehensive classroom behavior assessment. For school-age children with developmental disabilities, classroom behavior management and interventions are the primary methods for improving their classroom performance. There are three common primary types of classroom behavior interventions: one-on-one peer help or parent coaching [18], instructional task modification [19], and self-monitoring [20]. However, while this traditional intervention method has helped children with developmental disabilities improve their classroom task completion rate and classroom behavioral performance, this behavioral intervention undoubtedly consumes many resources in terms of monitoring children's classroom behavior. It requires significant human and material resources to assist children's classroom learning process.

1.2.2. Advanced Methods on Human Activity Recognition

With the development and popularization of artificial intelligence technologies, research on scenario-based understanding and behavioral analysis has shone in practical application scenarios [21]. With the help of data-driven and algorithmic reasoning, machine learning theory provides the feasibility of achieving a one-to-one accurate understanding and assessment of students' classroom behaviors [22], especially for fine-grained behavioral analysis that is difficult to be taken into account by manual statistics. Some scholars have already implemented AI techniques with classroom scene understanding and achieved better results. For example, intelligent classroom systems that assist teachers in teaching and personalize students' learning by building front-end interactive learning environments for teachers and students and back-end intelligent learning management systems [23]. Adaptive education platforms that solve students' specific learning problems provide personalized teaching and improve students' learning experiences according to their needs and their abilities [24]. However, little research has been conducted on AI-based classroom behavioral understanding due to the immature combination of technologies and the niche nature of the educational scenario, making it almost impossible to find corresponding work for reference. Although classroom behavioral activities are too complex and refined, it still belongs to the domain of behavior recognition, so we can help build an intelligent classroom-acceptable behavior perception system by referring to the relevant theories of human activity recognition. A brief overview of the human activity recognition approach is presented in the following section.

The mainstream approaches for human activity recognition can be roughly classified into vision-based and sensor-based based on different data sources [25]. Vision-based behavioral analysis systems usually use single or multiple RGB or depth cameras to collect images or video data of participants' behavioral information, environment, and background information in a specific activity space [26]. Moreover, after feature extraction of the collected data through image processing techniques and computer vision methods, they can be used to identify participants' behavior through algorithm learning and inference. Research conducted by numerous scholars applying vision methods in the field of human activity recognition includes: identifying group behavior and classifying abnormal activities in crowded scenes for surveillance as well as public personal safety purposes [27–29],

including analysis of fall detection, patient monitoring and other behavioral recognition of individuals to improve the quality of human life through vision [30,31]. However, since the vision-based data acquisition equipment for human activity behavior recognition mainly relies on cameras, it is vulnerable to environmental conditions such as light and weather, the shooting range and angle, and a large amount of acquired data storage. The reference of participants' activity is easily affected by environmental occlusion and privacy issues. Due to the influence of these factors, vision-based behavior analysis systems have not yet been widely used. In contrast, sensors have the advantages of high sensitivity, small size, accessible storage, and wide applicability to various scenarios, which can avoid various problems in using vision devices, so they are now widely embedded in mobile phones, smartwatches and bracelets, eye-tracking devices, virtual/augmented reality headsets, and various intelligent IoT devices [32]. Meanwhile, along with the widespread popularity of mobile Internet and the increasing demand for daily public use of intelligent devices, the problems of inconvenience in carrying and limited endurance of traditional sensor-based devices have been effectively solved in various application scenarios, and they have now become one of the mainstream methods for human activity recognition [25]. Scholars have applied sensing devices for intelligent activity recognition in several daily domains: Alani et al. achieved 93.67% accuracy in 2020 using a deep learning approach to recognize twenty everyday human activities in intelligent homes [33]; Kavuncuoğlu et al. used only a waist sensor to achieve accurate monitoring of fall and daily motion data, achieving 99.96% accuracy in 2520 data [34].

1.3. Contributions and Structure

The contributions of this paper include: 1. Artificial intelligent based behavior recognition is applied to the classroom environment for the first time, and an intelligent system with motion sensors to perceive and identify classroom behavior is built. 2. Based on sensor hardware devices, a classroom behavior database (SCB-13) including 14 common classroom behaviors collected from 13 participants is constructed. 3. A method of extracting valid sensor data segments based on an improved Voting-Based Dynamic Time Warping algorithm (VB-DTW) is proposed. 4. An intelligent identification method is proposed to recognize 14 common classroom behaviors based on valid behavior segments combined with a 1DCNN algorithm, and the proposed method achieved 100% recognition accuracy on a self-constructed dataset (SCB-13).

The second part of this paper describes the data hardware acquisition system and the relevant characteristics of the data; the third part gives a brief overview of the basic principles of the algorithm; the fourth part is the experimental results and comparative analysis; finally, the paper provides a conclusion.

2. Materials and Methods

2.1. Participants

In this study, we recruited 13 participants to carry out a feasibility study on the possibility of accurately identifying students' classroom behavior. The participants, aged from 20 to 26 years, were invited to participate in a classroom behavioral simulation experiment. This population consisted of 6 males and 7 females without special educational needs or developmental problems. They were culturally literate and able to comprehend, imitate, and model classroom behaviors accurately. Participants signed consent forms approved by the Ethics Committee of The Education University of Hong Kong (Approval Number: 2021-2022-0417) before data collection.

2.2. Experimental Design

For each participant, 5 sets of experimental data were gathered, and a total of 65 sets of data were collected. In each trial, participants were tasked with simulating 14 common classroom behaviors, and Table 1 shows the design of each motion. Each motion lasts for 20 s, which can be divided into the valid time duration doing the motions and the sitting

still time. Except for when motion happens, the rest of the period is referred to as sitting still time.

Table 1. Motion mode design. To simulate classroom behaviors for the participants, we selected 14 typical classroom behaviors. The table lists each motion's name as well as the order in which it took place.

Serial No.	Motion Mode
1	Sitting still
2	Lying on the desktop
3	Writing notes
4	Raising a hand in the seat
5	Turning around and looking around
6	Raising a hand while standing up
7	Rocking on the seat
8	Standing up and sitting down
9	Wandering and trunk rotation
10	Playing hands
11	Turning pen in hand
12	Knocking on the desktop
13	Leaning the body and chatting
14	Shaking legs

The actual hardware system used in the acquisition system is MPU6050, the main hardware processing chip is ESP-8266, the acquisition data bit rate is 115,200 Hz, the Arduino hardware platform is used for programming control, and the sensor data is stored in a .CSV file format via the computer's USB port using Python program. Figure 1a illustrates the schematic diagram of the 3D acquisition system, and 3 cameras are respectively installed on the participant's left side, front side, and diagonal rear to record the participant's vision motion data. As depicted in Figure 1b, in order to investigate the effect of the sensor in different positions, the sensors are positioned in the middle of the spine and the right shoulder of the participant. In addition to the 14 motions, there is a 20-s system calibration time at the beginning of the experiment to reduce the initial error caused during data acquisition, and the total duration of each experiment is 5 min. The sensors generate 7 channels of data: accelerometer (x -axis, y -axis, and z -axis) data, gyroscope (x -axis, y -axis, z -axis) data, and temperature data. The participants' motion information can be measured using accelerometer data in various directions. Gyroscope data can monitor angular velocity to determine an object's position and rotational orientation. Due to their susceptibility to environmental factors, temperature data are insufficient for use in a motion recognition system.

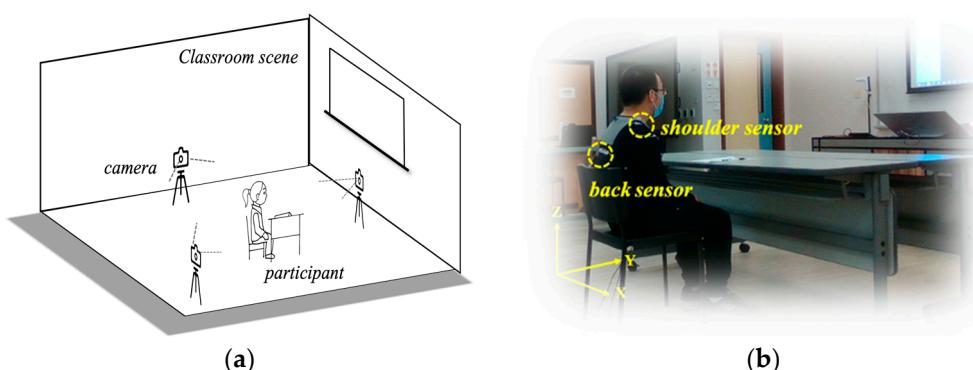


Figure 1. The acquisition system of the experiment. (a) The schematic diagram of the motion acquisition system in the classroom scene; (b) The location of the sensors. The vision information of the participants' motions is collected through cameras from three perspectives to assist in the classification. One sensor was placed in the center of the participant's spine and another one on the right shoulder to collect data on the participant's motions.

2.3. Experiment Data Introduction

SCB-13: The self-built dataset SCB-13 of this paper is made up of the above 13 participants' classroom behavior sensor data. The dataset will be used for later data analysis and model accuracy testing. Furthermore, we intend to provide a brief explanation of the experiment data of the back sensor from the perspective of an intuitive explanation.

2.3.1. Multiple Channel Data Display

After separating the gathered data by 14 given motion patterns, the 65 sets of data for the same motion are averaged to eliminate individual motion differences. Figure 2 displays the processed 6-channel data when motion 6 (raising hand while standing up) is selected as a sample motion. The data demonstrate that motion occurrence and stable state can be acquired during the valid duration of motion and sitting still time, respectively. Notably, when the students get up and raise their hands, the Z-axis data of the accelerometer change the most, which is consistent with the actual situation. It confirms the viability of using sensors for behavior identification.

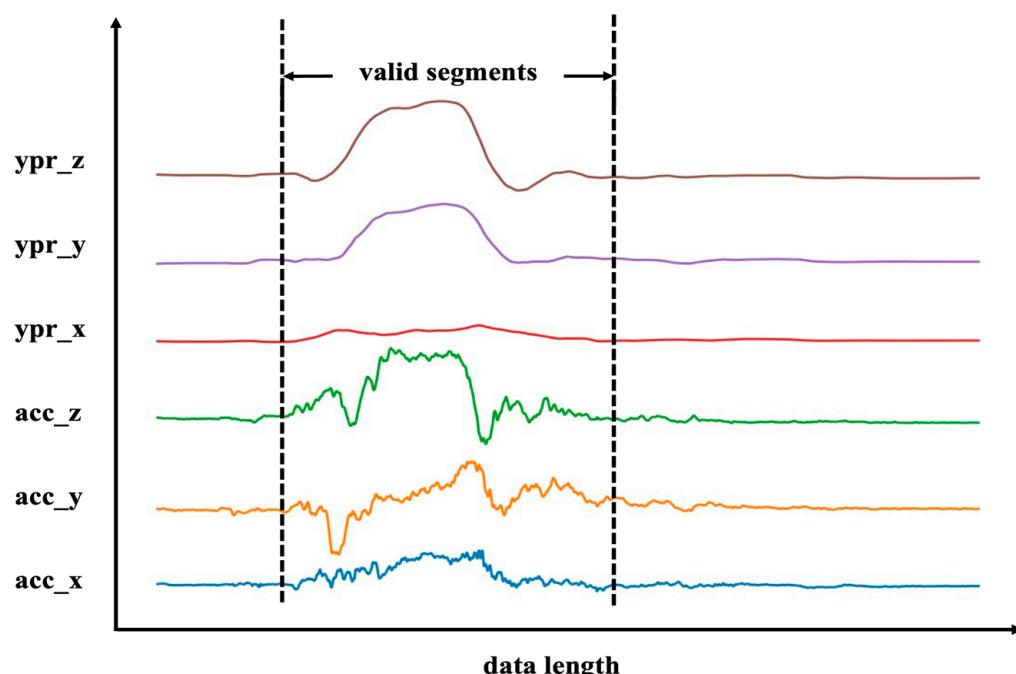


Figure 2. Take action 6 as an example to display the data of each channel of the back sensor. The lines from bottom to top represent the accelerometer x -axis(acc_x), y -axis(acc_y), z -axis(acc_z), gyroscope x -axis(ypr_x), y -axis(ypr_y), and z -axis(ypr_z). Valid segments of motions are shown within dashed lines.

2.3.2. Display of Different Motions of the Same Participant

A participant was selected randomly, and his/her 4 common classroom behaviors (motion 1 sitting still, motion 5 turning around and looking around, motion 6 raising hand while standing up, and motion 8 standing up and sitting down) were displayed in the accelerometer(acc) channel and the gyroscope(ypr) channel, as shown in Figure 3. There are observable changes in the data between different motions of the same volunteer. Nonetheless, the data patterns of motions 6 and 8 are comparable to some extent, providing the classification a challenging problem.

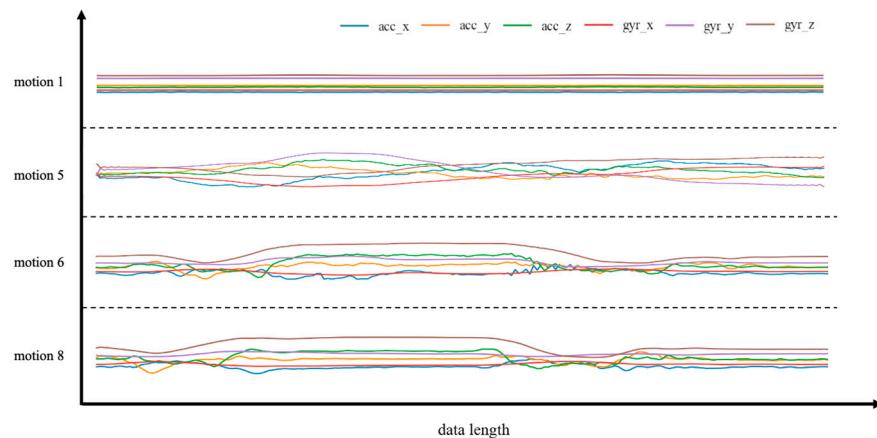


Figure 3. The randomly-selected four different actions of one of the participants and the data of the accelerometer and gyroscope data of the back sensor. The selected actions are as follows: motion 1 (sitting still), motion 5 (turning around and looking around), motion 6 (raising hand while standing up) and motion 8 (standing up and sitting down). We uniformly downsampled the data length to 200 for display clarity. Through motion sensors, we can continuously collect data about different motions, and each motion has a unique motion pattern. The relative intensity of each action is reflected in the ordinate after normalization.

2.3.3. Display of Different Participants with the Same Motion

Figure 4 shows the accelerometer data and gyroscope data of motion 6, raising hands while standing up, which were collected from 4 randomly picked individuals in order to display the differences in motion between various participants.

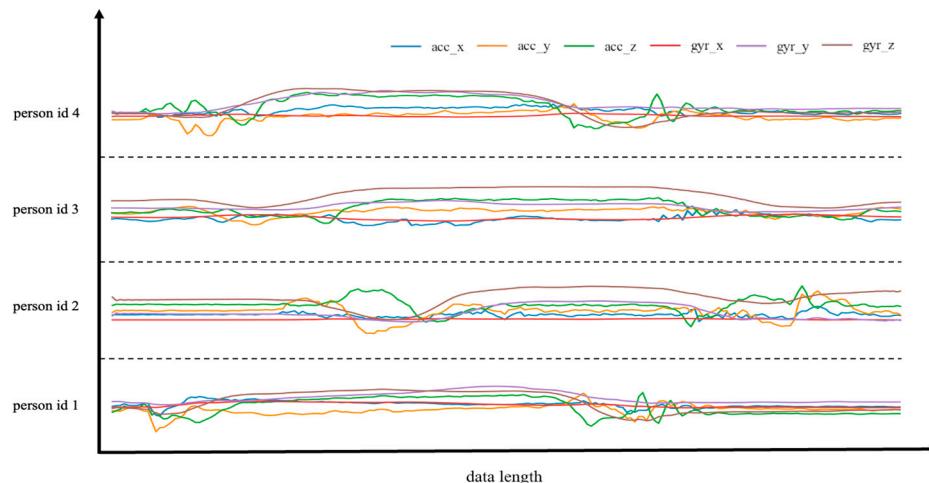


Figure 4. In the same action mode, the data of four volunteers (id1, id2, id3, id4) are randomly selected for display. We uniformly downsampled the data length to 200 for display clarity. It was challenging to classify classroom behavior since each participant carried out the same action in different ways and had unique sensor data patterns.

The preceding diagram demonstrates that various participants have distinct motion pattern characteristics, even for identical motion. It may be caused by variances in personal posture and habitual behaviors. This necessitates that the established model has robust generalization performance, capable of identifying the distinctions between the characteristics of various motion patterns while allowing for modest variations within the same motion. A comparison of the accelerometer and gyroscope data determined that the gyroscope data has more complex properties and fewer noise points, making it more ideal for the learning and reasoning of the neural network. Before generating the network's stan-

dard input, it is necessary to address the extraction and separation of valid data segments since the same motion of different participants occurs at different times and lasts varying times. Taking into account the temporal features of the data, we attempt to extract valid segments of the entire motion time in this article, which was detailed demonstrated in the identification algorithm.

2.4. Identification Algorithm

Overall, the algorithm is divided into 3 stages: the extraction of valid segments based on the Dynamic Time Warping algorithm, data augmentation, and a deep learning-based classification algorithm. The whole process of the algorithm can be shown in Figure 5. Further, about the classification algorithm, we picked the most typical Deep Neural Networks (DNNs) as the classification benchmark and investigated the classification accuracy of the RNN-based method and the CNN-based method to explore the impact of various algorithms on the precise perception and identification of classroom behaviors.

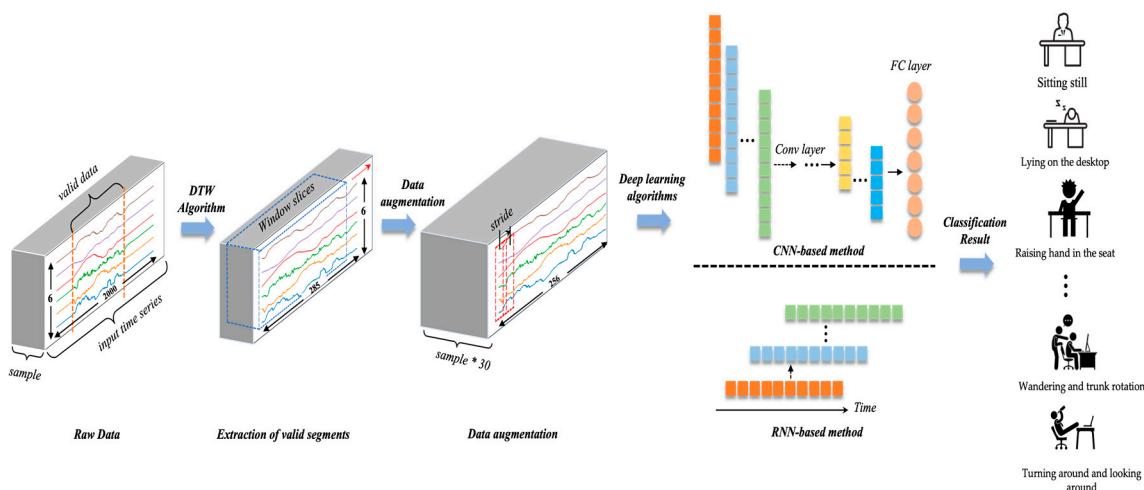


Figure 5. The framework of the whole process of the algorithm. The algorithm takes raw data as the input and outputs the most likely behavior from the 14 common classroom behaviors.

2.4.1. Voting-Based DTW (VB-DTW) Valid Segment Extraction Algorithm

Initially, we normalized the collected data to eliminate large differences in data values, which can hinder the convergence of the model. We scale the features contained in each channel by the maximum value and minimum value to the interval $[0, 1]$ without affecting the numerical distribution:

$$X = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Developing a distinctive and suitable method for feature representation is necessary in order to assess if motions can be accurately distinguished from the continuous and substantial stream of sensor data. The classification accuracy is determined by the algorithm's capacity to accurately extract the features in each motion sequence, particularly for sequences having temporal properties. Even though each motion's recommended acquisition time is equal, the valid duration of each motion varies due to participant differences during the acquisition process. The ratio of the motion's valid segment to its total time segment is insufficient for some motions (such as raising a hand on a seat, standing up and raising a hand, standing up and sitting down, and knocking on a table), making it challenging to identify motion patterns and represent motion features. To accurately identify the motion mode of each motion, we must differentiate the sitting still state from the valid duration data. In this context, we proposed an improved algorithm for signal extraction based on the Dynamic Time Warping (DTW) algorithm [35], which names as Voting-Based DTW (VB-DTW) valid segment extraction algorithm.

Since the valid motion segments are surrounded by the “sitting still” data in this work, we must divide the raw motion data into tiny sequences to efficiently locate the valid segment rather than process an entire motion segment directly. To extract the valid segments, we divide the raw motion data with a length of 50, which splits the entire 2000 motion data into 40 smaller slices. Utilizing the VB-DTW algorithm, we figure out the minimum warped path of 2 adjacent slices, a total of 39 warped path values yields for each motion from a total of 40 slices. The average warped path value of the motion is utilized as the threshold, and the combined vote of 4 neighboring warped paths is used to evaluate if the slices correspond to valid motion clips. The effective segment length of the final motion is determined by connecting the extracted valid segments. In addition, to address the issue of varied lengths for each extracted valid motion segment, we uniformly downsample the extracted valid motion segments to 285 in order to make model training easier. We apply the VB-DTW algorithm on the remaining thirteen types of motions except the sitting still since the complete motion sequence of the sitting still is a valid segment of the motion. As a result, we directly downsample the sitting still data to a length of 285. The whole process of the VB-DTW-extracted valid segment algorithm can be shown in Algorithm 1.

Algorithm 1: Voting-Based DTW (VB-DTW) valid segment extraction algorithm

Input:

Segments of original data : $S_i, i = 1, 2, 3, \dots, N$, where $N = 40$;
For each time sequence : $S_i = S_i^1, S_i^2, S_i^3, \dots, S_i^M$, where $M = 50$.

Initialization:

```

 $Dv = \{ \}$ ,  $Sv = \{ \}$ 
voting set =  $\{ D_{j-2}, D_{j-1}, D_j, D_{j+1}, D_{j+2} \}$ 
1: while  $1 \leq i \leq N - 1$  do
2:    $D_i \leftarrow DTW(S_i, S_{i+1})$ 
3: end while
4:  $threshold \leftarrow \frac{1}{N-1} \sum_{i=1}^{N-1} D_i$ 
5: while  $3 \leq j \leq N - 2$  do
6:   for  $D_{value}$  in voting set do
7:     count  $\leftarrow 0$ 
8:     if  $D_{value} > threshold$  then
9:       count  $\leftarrow count + 1$ 
10:    if count  $\geq 3$  then
11:       $Dv \leftarrow Dv \cup j$ 
12:     $j; j \leftarrow j + 1$ 
13: end while
14: for  $j$  in  $\{1, 2, N - 1, N\}$  do
15:   if  $D_j >$  then
16:      $Dv \leftarrow Dv \cup j$ 
17:  $Sv \leftarrow Sv \cup \{ S_{D_v^k}, S_{D_v^k+1} \}$  ( $k = 1, 2, 3, \dots, length(Dv)$ )

```

Output:

DTW value of time series : $D_j, j = 1, 2, 3, \dots, N - 1$;
Valid segment slices : $S_v, S_v \in S_i$.

2.4.2. Data Augmentation

Data augmentation assists in resolving the overfitting issue caused by insufficient data sets during model training. Contrary to the data augmentation methods for image data, time series data augmentation confronts several formidable obstacles, including 1. the fundamental features of time series sequences are underutilized, 2. different jobs necessitate the use of distinct data augmentation techniques, and 3. the issue of sample category imbalance.

Traditional time series data augmentation methods can be subdivided into time domain-based data enhancement to convert original data or to inject noise; frequency domain-based data enhancement converts data from the time domain to the frequency

domain and then applies enhancement algorithms; and simultaneous time domain and frequency domain analysis. To prevent the issue of model overfitting caused by insufficient data, to strengthen the model's robustness, and to generate a high number of data samples, we use the window slicing-based method as the data enhancement technique. Window slicing separates the original data of length n into $n - s + 1$ slices with the same label as the raw segment, using S as the new slice length. During the training process, each slice is sent to the network independently as a training instance for prediction. During testing, the separated slices are also submitted to the network, and the majority vote is utilized to determine the original segment's label. In this model, we select a slice length of 256, which corresponds to approximately 90% of the original length of 285. Figure 6 depicts the data augmentation method, which divides the down-sampled valid motion sequence into 30 new slices.

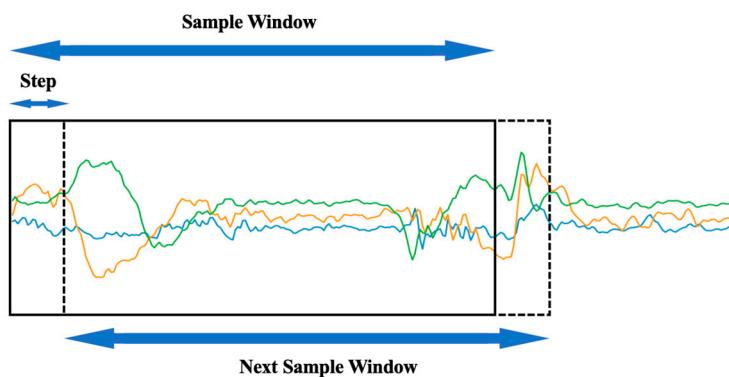


Figure 6. The detailed data augmentation procedure. The green/orange/blue lines represent the sensor data for a motion. The detailed data augmentation procedure. The green/orange/blue lines represent the sensor data for a motion. The sample window size is 256, and the stride size of the window is 1. We received 30 identical labeled data for each 285-length motion data after data augmentation.

2.4.3. Deep Learning-Based Classification Algorithm

We explored 2 categories, Recurrent Neural Networks (RNN) based methods and Convolutional Neural Network (CNN) based methods. The RNN-based method tries to represent data attributes based on temporal properties. Long Short Term Memory network (LSTM) [36] and Bidirectional Long Short Term Memory network (BiLSTM) [37] are the specific algorithms chosen for RNN-based methods. The CNN-based method can extract features by performing convolution on the data and focusing on the data's spatial characteristics. The chosen method for CNN-based methods is 1DCNN [38]. The basic deep neural network (DNN) is chosen as a simple benchmark model that aims to evaluate the performance of various algorithms from these 2 categories. The reason for comparing these four models is that this paper aimed to explore the more classical, advanced, and effective models of temporal data processing for the performance of perception and identification of students' classroom behavior tasks. The choice of these four classical models helped us to achieve the goal of presenting the best results of our model compared to the rest of the models.

(1) LSTM and BiLSTM

Recurrent neural networks (RNNs) are uniquely valuable compared to other neural networks for processing interdependent sequential data, such as text analysis, speech recognition, and machine translation. It is also widely used in the field of sensor-based motion recognition due to its property of recursion in the direction of sequence evolution, and all recurrent units are linked in a chain [39].

However, the conventional RNN has a short-term memory problem because the RNN cannot memorize and process more comprehensive sequence information, as the layers

in the pre-recursive stage will stop learning due to the vanishing gradient problem or exploding gradient problem caused by backpropagation. For the problem that the later data input has more influence and the earlier data input has less influence on RNN, in 1997, Hochreiter and Schmidhuber proposed the Long Short Term Memory Network (LSTM), which successfully solved the limitation of RNN in processing long sequence data and was able to learn the long-term dependence of sequence data features. LSTM proposed the internal mechanism of ‘gates’ used to regulate the flow of feature information, including input gates that control the reading of data into the unit, output gates that control the output entries of the unit, and forgetting gates that reset the contents of the unit. The specific LSTM structure is shown in Figure 7, and a new vector C representing the cell state is added to the LSTM.

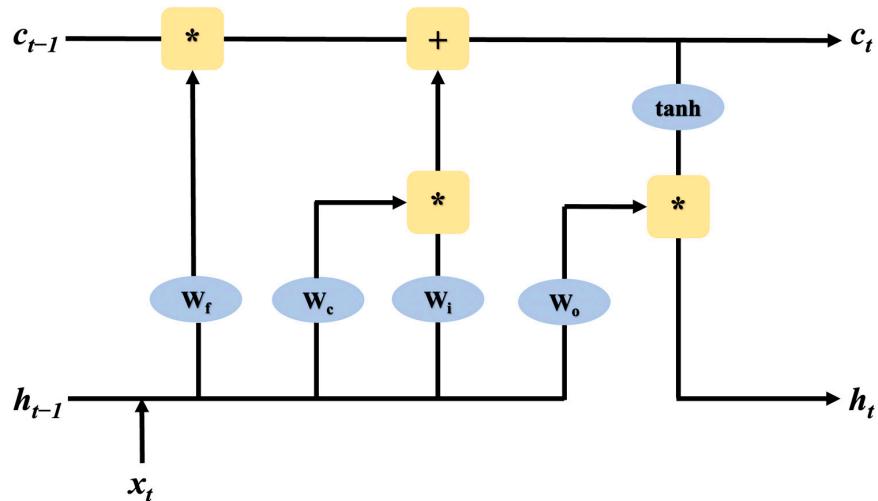


Figure 7. LSTM structure, W_f is the forgetting gate, W_i is the input gate, W_o is the output gate, x_t is the input data, h_{t-1} is the neural node of the hidden state, and W_f is used to calculate the features in c_{t-1} to obtain c_t .

Both traditional RNN and LSTM can only predict the output of the next moment based on the information of the previous moment. While in practical applications, the information of the next moment may also have a significant influence on the output state of this moment. Bi-directional LSTM (Bi-LSTM) combines 2 traditional LSTM models and uses 1 of them for forward input and the other for reverse input to fuse the information of the previous and subsequent moments for inference. Its structure is shown in Figure 8.

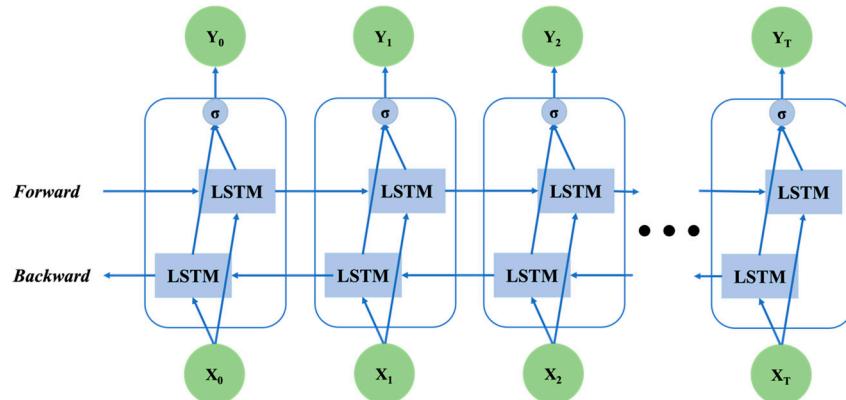


Figure 8. Bi-LSTM structure, which combines forward LSTM and backward LSTM.

One-dimensional convolutional neural networks (1DCNN) have strong advantages for sequence data because of the powerful ability to extract features from fixed-length segments in 1-dimensional signals. Also, the adaptive 1DCNN only performs linear 1D convolutions (scalar multiplication and addition), thus providing the possibility of real-time and low-cost intelligent control over hardware [40]. The basic structure of 1DCNN is shown in Figure 9. The kernel moves on the sequence data along the time axis to complete the feature extraction of the original data.

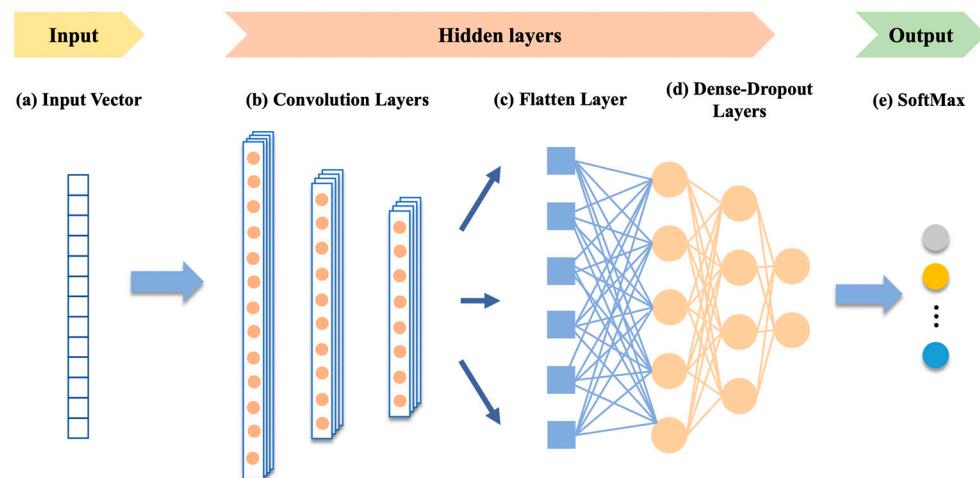


Figure 9. 1DCNN structure. The structure of 1DCNN mainly includes input, hidden layer, and output, so as to achieve the purpose of feature extraction.

In conclusion, the algorithm utilized the VB-DTW algorithm to extract valid segments, and then window slicing was used to augment the data and achieve a 30-times dataset increase. For classification, we employ 2 categories of networks. For the RNN-based method, the LSTM network and Bi-LSTM network are chosen, as well as the 1DCNN for the CNN-based method. These 2 different types of networks' abilities and contributions to percept and identify students' classroom behavior are assessed.

2.4.4. Evaluation Metrics

(1) Valid Segments Extraction

In order to demonstrate the accuracy of the valid segments obtained by the VB-DTW algorithm, we hand-crafted labeled the indices of all valid motion segments as the benchmark. We measure the similarity between the index of extracted data slices (represented as A) and the benchmark (represented as B) using the Jaccard index. The Jaccard index is used to determine the degree of similarity between limited sample data and is defined as the sample intersection size divided by the sample union size. The equation is:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (2)$$

(2) Motion Identification

In order to verify the classification performance of the model, we usually use the accuracy rate to characterize it, that is, the proportion of the number of samples with accurate classification (represented as a) to the total number of samples (represented as m) of this type. Expressed by the following formula:

$$\text{accuracy} = \frac{a}{m} \quad (3)$$

3. Results

In summary, based on the need to understand the classroom behaviors of school children in educational scenarios, sensor-based devices provide an effective way to identify classroom behaviors intelligently. Therefore, this paper proposes the VB-DTW algorithm based on wearable sensors combined with artificial intelligence technology to achieve intelligent recognition of school children's classroom behaviors. Based on the recognition results, it is possible to provide immediate feedback on students' classroom performance and help them improve their learning performance while providing an essential reference basis and data support for constructing an intelligent digital education platform.

3.1. Identification Algorithm Valid Segmentation Results

For the 65 groups of motions with the same label, we calculate the Jaccard index of each channel of acc and ypr and then determine the average Jaccard index for each motion by averaging the six-channel values. As shown in Table 2, all the extracted valid segments' indices except lying on the desktop and writing notes are more than 88% similar to the benchmark. The Jaccard index of lying on the desktop and writing notes is worse than other motions, which may be due to the sensor data not changing significantly during motion times, as well as the warped path between the adjacent paths being near. This is a weakness in our proposed VB-DTW algorithm, which makes the algorithm inefficient for long-term recognition of a substantial portion of near-static data. We will continue to investigate the most effective approach to dealing with precise and effective segment extraction in subsequent tests.

Table 2. Jaccard index for 13 motions. All the extracted valid segments' indices except lying on the desktop and writing notes are more than 88% similar to the benchmark.

Motion Mode	Jaccard Index
Raising a hand in the seat	0.97
Turning around and looking around	0.96
Raising hand while standing up	0.96
Rocking on the seat	0.97
Stand up and sit down	0.97
Wandering and trunk rotation	0.98
Playing hands	0.87
Turning pen in hand	0.88
Knocking on the desktop	0.95
Leaning the body and chat	0.96
Shaking legs	0.94
Lying on the desktop	0.45
Writing notes	0.50

3.2. Motion Identification Results

Furthermore, the performance of the aforementioned four models in accurately classifying classroom behavior is evaluated in order to measure the influence of different classification models on the self-constructed dataset (SCB-13). A deep neural network (DNN) is chosen as a simple benchmark model for the purpose of evaluating the efficacy of various algorithms. Separately for the back sensor and shoulder sensor, the research tests the accelerometer data (acc), gyroscope data (ypr), and accelerometer and gyroscope data (acc + ypr). The research confirms the effect of classifying sensor data using LSTM and BiLSTM networks, respectively, taking into account the time-series characteristic of the data. In addition, from the perspective of one-dimensional signal feature extraction, the research uses 1DCNN to extract and classify data features in a more "intelligent" mode. The results of the experiments carried out are listed in Table 3 below.

Table 3. Main Result of four networks for the back sensor and shoulder sensor separately. Furthermore, acc represents accelerometer data, ypr represents gyroscope data, and acc + ypr represents the combination of accelerometer and gyroscope data.

Accuracy (%)	Back			Shoulder		
	acc	ypr	acc + ypr	acc	ypr	acc + ypr
DNN	81.8	91.2	93.3	89.5	86.5	91.7
LSTM	66.5	84	96.4	81.3	81.6	89.2
BiLSTM	96	98	99.8	96.4	95.9	97.2
1DCNN	99.8	99.9	100	99.6	98.3	98.8

Based on a comprehensive evaluation of the experiment outcomes, we have determined that both DNN and LSTM networks are generally useful in distinguishing classroom behaviors from the three channels' data of the accelerometer or gyroscope. However, when accelerometer and gyroscope data are incorporated into the network input, the classification effect of the DNN and LSTM network is significantly enhanced, demonstrating that more data channels are beneficial for the expression and differentiation of features.

The main experiment results show that, compared to DNN and LSTM networks, the BiLSTM network significantly improves the identification accuracy of classroom behavior. In addition, BiLSTM networks are capable of a more robust feature representation, whether for three-channel data (accelerometer, gyroscope) or six-channel data (accelerometer and gyroscope), demonstrating that the combination of forward-backward LSTM neural network for the learning of feature representation has been significantly improved.

Compared to the other three networks, the unique and potent feature extraction capabilities for sequence data demonstrated by the 1DCNN network stands out. Combining accelerometer and gyroscope data, the 1DCNN achieves classification accuracy of 100% and 98.8% for the back and shoulder sensors, respectively. In terms of model complexity and computing speed, 1DCNN is considerably superior to LSTM and BiLSTM.

In general, the data collected by the back sensor is more stable than that collected by the shoulder sensor, allowing for the differentiation of classroom activities on a wider scale. For motion classification, the gyroscope is superior to the accelerometer, despite neither being as accurate as when accelerometer and gyroscope data are used simultaneously in the classification.

4. Discussion

4.1. Ablation Study

4.1.1. Effect of VB-DTW Valid Segment Extraction

To evaluate the effectiveness of the proposed VB-DTW algorithm for valid segment extraction, we chose the data with the best classification impact (the combination of acc and ypr data) to investigate how valid segment extraction affected the action classification results. Table 4 displays the test results. According to the test results, it can be inferred that the results with VB-DTW valid segment extraction generally have higher accuracy than those without VB-DTW. The 1DCNN model outperforms the other algorithms in terms of classification accuracy for valid segment extraction.

Table 4. Test result of the effectiveness of VB-DTW valid segment extraction.

Accuracy (%)	With VB-DTW Valid Segment Extraction		Without VB-DTW Valid Segment Extraction		Improvement by VB-DTW	
	Back	Shoulder	Back	Shoulder	Back	Shoulder
DNN	93.3	91.7	89.6	82.5	3.7↑	9.2↑
LSTM	96.4	89.2	92.1	85.2	4.3↑	4.0↑
BiLSTM	99.8	97.2	93.8	90.2	5.0↑	7.0↑
1DCNN	100	98.8	98.5	95.9	1.5↑	2.9↑

4.1.2. Effect of VB-DTW Augmentation

In order to compare the accuracy of the model with and without data augmentation, we still select the data (the combination of Acc and Ypr data) with the highest level of classification accuracy. Table 5 displays the test results. The test results show that the model's classification accuracy with and without data augmentation is significantly different, and the special benefits of 1DCNN in the categorization of time series data are not reflected. These results might be brought on by the issue of data overfitting by the insufficient amount of data we gathered. As a result, for datasets with fewer data, the proposed algorithm needs to apply data augmentation on the dataset.

Table 5. Test result of the effectiveness of data augmentation.

Accuracy (%)	With VB-DTW Augmentation		Without VB-DTW Augmentation		Improvement by VB-DTW	
	Back	Shoulder	Back	Shoulder	Back	Shoulder
DNN	93.3	91.7	41.2	39.0	52.1↑	52.7↑
LSTM	96.4	89.2	49.5	51.1	46.9↑	38.1↑
BiLSTM	99.8	97.2	52.8	51.1	47.0↑	46.1↑
1DCNN	100	98.8	53.8	52.1	46.2↑	46.7↑

According to the results and discussions, the proposed VB-DTW algorithm, based on wearable sensors and artificial intelligence technology, achieves intelligent perception and identification of school-aged students' classroom behaviors. Furthermore, effective, valid segment extraction methods, as well as data augmentation in model design, are essential for the network's superior performance. Intelligent recognition of school-age children's classroom behavior can provide timely feedback, allowing the children, particularly those with special education needs, to grasp their classroom behavior in real-time and obtain assistance in the classroom without being labor-intensive.

4.2. Limitation of the Proposed Method

However, the proposed method has several limitations, particularly when students' classroom behaviors do not change significantly over time (e.g., writing notes). The proposed method cannot efficiently extract the segments of students' motions. This issue happened because the segments could not be extracted successfully due to the warped path of the DTW algorithm between adjacent paths being near since the absence of significant changes in the sensor data during motion. As a result, the proposed VB-DTW algorithm is inefficient for the long-term recognition of the majority of near-static data. In future work, we will still explore the most efficient way of dealing with precise and valid segment extraction.

5. Conclusions

The purpose of this paper is to provide auxiliary education by intelligently perceiving the behavior of students during classroom scenarios by integrating sensor equipment with AI technology. In this article, an improved algorithm which was named VB-DTW is proposed for separating valid sensor signals based on the DTW algorithm, and the effectiveness is validated using the Jaccard index. It provides the capacity to discern accurately between static and dynamic data. In addition, four classical deep learning network structures are compared for the accuracy of classroom behavior classification. It is discovered that the 1DCNN algorithm has the highest accuracy rate, particularly when accelerometer and gyroscope data are aggregated, where the recognition accuracy rate reaches 100%. We anticipate classifying more classroom activities based on hardware in real time and achieving multi-modal identification by fusing sensor data and visual data in future studies.

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

Funding: This research was supported a Dean’s Research Fund (2021/22 DRF/SRAS-1/9th), the Education University of Hong Kong. This research was also supported by Wuxi Taihu Lake Talent Plan Supporting for Leading Talents in Medical and Health Profession, China.

Institutional Review Board Statement: The ethical review board at the Education University of Hong Kong approved this study (Protocol code: 2021-2022-0417).

Informed Consent Statement: Written informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to acknowledge The Education University of Hong Kong for the support in the provision of experimental sites.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhu, Z.-T.; Yu, M.-H.; Riezebos, P.J.S.l.e. A research framework of smart education. *Smart Learn. Environ.* **2016**, *3*, 4. [[CrossRef](#)]
2. Shoikova, E.; Nikolov, R.; Kovatcheva, E. Smart digital education enhanced by AR and IoT data. In Proceedings of the 12th International Technology, Education and Development Conference (INTED), Valencia, Spain, 5–7 March 2018; pp. 5–7.
3. Atabekov, A. Internet of things-based smart classroom environment: Student research abstract. In Proceedings of the 31st Annual ACM Symposium on Applied Computing, 2016, Pisa Italy, 4–8 April 2016, pp. 746–747.
4. Zhan, Z.; Wu, Q.; Lin, Z.; Cai, J. Smart classroom environments affect teacher-student interaction: Evidence from a behavioural sequence analysis. *Australas. J. Educ. Technol.* **2021**, *37*, 96–109. [[CrossRef](#)]
5. Alghamdi, A.; Karpinski, A.C.; Lepp, A.; Barkley, J. Online and face-to-face classroom multitasking and academic performance: Moderated mediation with self-efficacy for self-regulated learning and gender. *Comput. Hum. Behav.* **2020**, *102*, 214–222. [[CrossRef](#)]
6. Brandmiller, C.; Dumont, H.; Becker, M.J.C.E.P. Teacher perceptions of learning motivation and classroom behavior: The role of student characteristics. *Contemp. Educ. Psychol.* **2020**, *63*, 101893. [[CrossRef](#)]
7. Khan, A.; Ghosh, S.K. Student performance analysis and prediction in classroom learning: A review of educational data mining studies. *Educ. Inf. Technol.* **2020**, *26*, 205–240. [[CrossRef](#)]
8. Hopman, J.A.B.; Tick, N.T.; van der Ende, J.; Wubbels, T.; Verhulst, F.C.; Maras, A.; Breeman, L.D.; van Lier, P.A.C. Special education teachers’ relationships with students and self-efficacy moderate associations between classroom-level disruptive behaviors and emotional exhaustion. *Teach. Teach. Educ.* **2018**, *75*, 21–30. [[CrossRef](#)]
9. Iadarola, S.; Shih, W.; Dean, M.; Blanch, E.; Harwood, R.; Hetherington, S.; Mandell, D.; Kasari, C.; Smith, T. Implementing a Manualized, Classroom Transition Intervention for Students with ASD in Underresourced Schools. *Behav. Modif.* **2018**, *42*, 126–147. [[CrossRef](#)] [[PubMed](#)]
10. Li, R.; Fu, H.; Zheng, Y.; Lo, W.L.; Yu, J.; Sit, H.P.; Chi, Z.R.; Song, Z.X.; Wen, D.S. Automated fine motor evaluation for developmental coordination disorder. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 963–973. [[CrossRef](#)]
11. Zablotsky, B.; Black, L.I.; Maenner, M.J.; Schieve, L.A.; Danielson, M.L.; Bitsko, R.H.; Blumberg, S.J.; Kogan, M.D.; Boyle, C.A. Prevalence and Trends of Developmental Disabilities among Children in the United States: 2009–2017. *Pediatrics* **2019**, *144*, e20190811. [[CrossRef](#)]
12. Johnson, K.A.; White, M.; Wong, P.S.; Murrihy, C. Aspects of attention and inhibitory control are associated with on-task classroom behaviour and behavioural assessments, by both teachers and parents, in children with high and low symptoms of ADHD. *Child Neuropsychol.* **2020**, *26*, 219–241. [[CrossRef](#)] [[PubMed](#)]
13. Dilmurod, R.; Fazliddin, A. Prospects for the introduction of artificial intelligence technologies in higher education. *ACADEMICIA Int. Multidiscip. Res. J.* **2021**, *11*, 929–934. [[CrossRef](#)]
14. Jo, J.; Park, K.; Lee, D.; Lim, H. An Integrated Teaching and Learning Assistance System Meeting Requirements for Smart Education. *Wirel. Pers. Commun.* **2014**, *79*, 2453–2467. [[CrossRef](#)]
15. Singh, H.; Miah, S.J. Smart education literature: A theoretical analysis. *Educ. Inf. Technol.* **2020**, *25*, 3299–3328. [[CrossRef](#)]

16. Lekwa, A.J.; Reddy, L.A.; Shernoff, E.S. Measuring teacher practices and student academic engagement: A convergent validity study. *Sch. Psychol. Q.* **2019**, *34*, 109–118. [[CrossRef](#)] [[PubMed](#)]
17. Porter, L. *Student Behaviour: Theory and Practice for Teachers*; Routledge: London, UK, 2020.
18. McMichan, L.; Gibson, A.M.; Rowe, D.A. Classroom-based physical activity and sedentary behavior interventions in adolescents: A systematic review and meta-analysis. *J. Phys. Act. Health* **2018**, *15*, 383–393. [[CrossRef](#)]
19. Cox, S.K.; Root, J.R. Modified Schema-Based Instruction to Develop Flexible Mathematics Problem-Solving Strategies for Students with Autism Spectrum Disorder. *Remedial Spec. Educ.* **2018**, *41*, 139–151. [[CrossRef](#)]
20. Bertel, L.B.; Nørlem, H.L.; Azari, M. Supporting Self-Efficacy in Children with ADHD through AI-supported Self-monitoring: Initial Findings from a Case Study on Tiimood. In Proceedings of the Adjunct 15th International Conference on Persuasive Technology, Aalborg, Denmark, 20–23 April 2020.
21. Kok, V.J.; Lim, M.K.; Chan, C.S. Crowd behavior analysis: A review where physics meets biology. *Neurocomputing* **2016**, *177*, 342–362. [[CrossRef](#)]
22. Zheng, R.; Jiang, F.; Shen, R. Intelligent student behavior analysis system for real classrooms. In Proceedings of the ICASSP 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 9244–9248.
23. Saini, M.K.; Goel, N. How Smart Are Smart Classrooms? A Review of Smart Classroom Technologies. *ACM Comput. Surv.* **2020**, *52*, 1–28. [[CrossRef](#)]
24. Kabudi, T.; Pappas, I.; Olsen, D.H.J.C.; Intelligence, E.A. AI-enabled adaptive learning systems: A systematic mapping of the literature. *Comput. Educ. Artif. Intell.* **2021**, *2*, 100017. [[CrossRef](#)]
25. Chen, K.; Zhang, D.; Yao, L.; Guo, B.; Yu, Z.; Liu, Y. Deep Learning for Sensor-based Human Activity Recognition. *ACM Comput. Surv.* **2022**, *54*, 1–40. [[CrossRef](#)]
26. Beddiar, D.R.; Nini, B.; Sabokrou, M.; Hadid, A. Vision-based human activity recognition: A survey. *Multimed. Tools Appl.* **2020**, *79*, 30509–30555. [[CrossRef](#)]
27. Bour, P.; Cribelier, E.; Argyriou, V. Crowd behavior analysis from fixed and moving cameras. In *Multimodal Behavior Analysis in the Wild*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 289–322.
28. Grant, J.M.; Flynn, P.J. Crowd Scene Understanding from Video. *ACM Trans. Multimed. Comput. Commun. Appl.* **2017**, *13*, 1–23. [[CrossRef](#)]
29. Sreenu, G.; Durai, S.J.J.o.B.D. Intelligent video surveillance: A review through deep learning techniques for crowd analysis. *Big Data* **2019**, *6*, 48. [[CrossRef](#)]
30. Nguyen, T.-H.-C.; Nebel, J.-C.; Florez-Revuelta, F.J.S. Recognition of activities of daily living with egocentric vision: A review. *Sensors* **2016**, *16*, 72. [[CrossRef](#)]
31. Prati, A.; Shan, C.; Wang, K.I.K. Sensors, vision and networks: From video surveillance to activity recognition and health monitoring. *J. Ambient. Intell. Smart Environ.* **2019**, *11*, 5–22.
32. Michail, K.; Deliparaschos, K.M.; Tzafestas, S.G.; Zolotas, A.C. AI-based actuator/sensor fault detection with low computational cost for industrial applications. *IEEE Trans. Control. Syst. Technol.* **2015**, *24*, 293–301. [[CrossRef](#)]
33. Alani, A.A.; Cosma, G.; Taherkhani, A. Classifying imbalanced multi-modal sensor data for human activity recognition in a smart home using deep learning. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), online, 19–24 July 2020; pp. 1–8.
34. Kavuncuoğlu, E.; Uzunhisarcıklı, E.; Barshan, B.; Özdemir, A.T.J.D.S.P. Investigating the Performance of Wearable Motion Sensors on recognizing falls and daily activities via machine learning. *Digit. Signal Process.* **2022**, *126*, 103365; [[CrossRef](#)]
35. Li, H.; Liu, J.; Yang, Z.; Liu, R.W.; Wu, K.; Wan, Y. Adaptively constrained dynamic time warping for time series classification and clustering. *Inf. Sci.* **2020**, *534*, 97–116. [[CrossRef](#)]
36. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
37. Zhou, P.; Shi, W.; Tian, J.; Qi, Z.; Li, B.; Hao, H.; Xu, B. Attention-based bidirectional long short-term memory networks for relation classification. In Proceedings of the Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers), Berlin, Germany, 7–12 August 2016; pp. 207–212.
38. Cho, H.; Yoon, S.M.J.S. Divide and conquer-based 1D CNN human activity recognition using test data sharpening. *Sensors* **2018**, *18*, 1055. [[CrossRef](#)] [[PubMed](#)]
39. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
40. Kiranyaz, S.; Ince, T.; Abdeljaber, O.; Avci, O.; Gabbouj, M. 1-D convolutional neural networks for signal processing applications. In Proceedings of the ICASSP 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, Great Britain, 12–17 May 2019; pp. 8360–8364.

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.