IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 23, 2019, accepted October 31, 2019, date of publication November 25, 2019,
date of current version December 9, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2955545

Wearable Sport Activity Classification Based on Deep Convolutional Neural Network

YU-LIANG HSU, (Member, IEEE), HSING-CHENG CHANG, AND YUNG-JUNG CHIU
Department of Automatic Control Engineering, Feng Chia University, Taichung 40724, Taiwan

Corresponding author: Yu-Liang Hsu (hsuyl@fcu.edu.tw)

This work was supported in part by the Ministry of Science and Technology of Taiwan, under Grant MOST 107-2221-E-035-080, MOST 108-2627-H-006-004, and MOST 108-2628-E-035-002-MY3.

ABSTRACT This paper develops a wearable sport activity classification system and its associated deep learning-based sport activity classification algorithm for accurately recognizing sport activities. The proposed wearable system used two wearable inertial sensing modules worn on athletes' wrist and ankle to collect sport motion signals and utilized a deep convolutional neural network (CNN) to extract the inherent features from the spectrograms of the short-term Fourier transform (STFT) of the sport motion signals. The wearable inertial sensing module is composed of a microcontroller, a triaxial accelerometer, a triaxial gyroscope, an RF wireless transmission module, and a power supply circuit. All ten participants wore the two wearable inertial sensing modules on their wrist and ankle to collect motion signals generated by sport activities. Subsequently, we developed a deep learning-based sport activity classification algorithm composed of sport motion signal collection, signal preprocessing, sport motion segmentation, signal normalization, spectrogram generation, image mergence/resizing, and CNN-based classification to recognize ten types of sport activities. The CNN classifier consisting of two convolutional layers, two pooling layers, a fully-connected layer, and a softmax layer can be used to divide the sport activities into table tennis, tennis, badminton, golf, batting baseball, shooting basketball, volleyball, dribbling basketball, running, and bicycling, respectively. Finally, the experimental results show that the proposed wearable sport activity classification system and its deep learning-based sport activity classification algorithm can recognize 10 sport activities with the classification rate of 99.30%.

INDEX TERMS Wearable inertial sensing device, sport activity classification, deep learning, convolutional neural network.

I. INTRODUCTION

With the rapid development of wearable intelligent and artificial intelligence technologies, performance analysis in sport science has undergone major changes in recent years. In general, manual recording and analysis performed by trained analysts in sport science has some disadvantages such as time intensive, time consuming, subjective in nature, and prone to human error and bias. Objective measurement and analysis for sport activities is essential to understanding the technical and physical demands associated with sports performance [1]. Automatic sport activity recognition (SAR) systems are developed to provide objective measurement and analysis in sport science, which have the potential to improve the accuracy and efficiency of sports performance analysis and evaluate the effectiveness of training programs designed by coaches. Common automatic SAR systems can be achieved through machine and deep learning approaches by using the data measured by computer vision and inertial sensing technologies [2]–[5].

Sport activities captured by computer vision can be utilized for athlete detection and tracking, pose estimation, tactical analysis, and movement analysis [6]. The procedures of the computer vision-based automatic SAR systems consist of player detection and tracking, temporal cropping, and targeted motion recognition, which are dependent on the sport type and camera setup [5]. The computer vision-based automatic SAR systems can provide rapid post-match analysis and real-time objective feedback before the next race for coaches and athletes. However, the computer vision-based automatic SAR systems suffer from limited environments, where the cameras are expensive, installed, and should

The associate editor coordinating the review of this manuscript and approving it for publication was Alicia Fornés.

capture all player in the installed environment, which influence the measurement and analysis performance.

The alternative to computer vision for sport activity detection is inertial sensing technology, where the sensors are wearable and composed of accelerometers, gyroscopes, and magnetometers. Wearable devices embedded with inertial sensors are widely used in numerous applications such as authentication, rehabilitation, gait analysis, health care, activity recognition, disease monitoring, navigation, and so on [7]–[12]. Recently, a number of researchers have developed wearable inertial-sensing-based sport activity monitoring and analysis systems for running, football, tennis, baseball, golf, fencing, basketball, cycling, badminton, table tennis, volleyball, and so on [3], [4], [13]–[23]. The inertial sensors can measure the acceleration, angular velocity, and magnetic signals generated by the sport motions during sport training and competitions, and further obtain the orientation, velocity, and trajectory of the sport travel through the integration process of the inertial signals. The advantages of the inertial-sensing-based automatic SAR systems include low cost, lightweight, small size, low power consumption, self-contained in operations, and being wireless.

In general, there are several signal processing procedures required to generate a suitable input data for machine learning classifier algorithms to recognize sport activities, which include filtering, window motion durations, signal normalization, feature extraction, feature reduction/selection, and classification. For example, Margarito *et al.* [24] selected the 11 acceleration features from the 13 most commonly time- and frequency-domain features by using the Relief method, which are input to the Naïve Bayes (NB), decision tree (DT), logistic regression (LR), and artificial neural network (ANN) classifiers for recognizing the 8 sport activities including cycling, cross trainer, rowing, running, squatting, stepping, walking, and weight lifting. Ermes *et al.* [25] utilized the custom decision tree combined with the ANN to classify the 9 sport activities which consist of lying down, sitting and standing, running, walking, rowing with a rowing machine, cycling with an exercise bike, Nordic walking, playing football, and cycling with a regular bike, using the 7 time- and frequency-domain features extracted from the accelerometers and GPS measurements. Mitchell *et al.* [26] presented the discrete wavelet transform (DWT)-based support vector machines (SVMs) optimized by the sequential minimal optimization (SMO) algorithm to recognize the 7 sport activities using the measurements of the smartphone accelerometers. Wang *et al.* [27] used the principle component analysis (PCA) to obtain 3 features from the 12 statistical features and 3 morphological features extracted from the microelectromechanical systems (MEMS) motion sensors' signals, which are input to the SVM classifier for classifying elite, sub-elite, and amateur volleyball players with an average accuracy of 94 %.

Deep learning techniques represent an alternative to these deep learning-based classifiers for dealing with inertial sensing data. Deep learning models can extract suitable discriminative features from inertial sensing data to train the deep learning models [22], [28]–[31]. In recent years, some researchers have focused on developing effective deep learning-based automatic SAR systems to measurements gathered from wearable inertial sensing devices. To name a few, Kautz *et al.* [32] employed a deep convolutional neural network (DCNN)-based sport monitoring system for beach volleyball using the Bosch BMA280 accelerometer sensor placed on the player's wrist. The proposed DCNN had successfully recognized the ten beach volleyball actions such as underhand serve, overhead serve, jump serve, underarm set, overhead set, shot attack, spike, block, dig, and null class. Jiao *et al.* [33] presented the DCNNs called GolfVanillaCNN, GolfVGG, GolfInception, and GolfResNet to classify the nine types of swing shapes using the strain-gage, accelerometer, and gyroscope sensors. The accuracy of the abovementioned DCNN-based classifiers reached 95.04%, 96.70%, 97.36%, and 92.07%, respectively.

From our literature review, many researchers in the field of machine learning have paid attention to the inertial-sensing-based sport activity classification tasks. However, there is a few of literature that utilizes deep convolutional neural networks (CNNs), which can extract inherent features from the inertial sensing measurements, to recognize sport activities. In this paper, a wearable sport activity classification system and its associated deep learning-based sport activity classification algorithm are presented for sport activity classification tasks. The wearable sport activity classification system consists of two wearable inertial sensing modules worn on athletes' wrist and ankle to collect motion signals of sport activities. The proposed deep learning-based sport activity classification algorithm composed of sport motion signal collection, signal preprocessing, sport motion segmentation, signal normalization, spectrogram generation, image mergence/resizing, and CNN-based classification is developed for classifying ten types of sport activities. The CNN classifier is composed of two convolutional layers, two pooling layers, a fully-connected layer, and a softmax layer, which is utilized to extract the inherent features from the spectrograms of the STFT of the inertial signals and divide the sport activities into table tennis, tennis, badminton, golf, batting baseball, shooting basketball, volleyball, dribbling basketball, running, and bicycling, respectively. The contribution of this paper is to develop a low cost wearable sport activity classification system and its deep learning-based sport activity classification algorithm for providing an effective tool for sport activity classification tasks, which utilizes the deep CNN classifier for obtaining more accuracy for sport activity classification.

The rest of this paper is organized as follows. In Section II, we introduce hardware architecture of the proposed wearable sport activity classification system in detail. The deep learning-based sport activity classification algorithm, consisting of sport motion signal collection, signal preprocessing, sport motion segmentation, signal normalization, spectrogram generation, image mergence/resizing, and CNN-based classification, is then described in Section III.

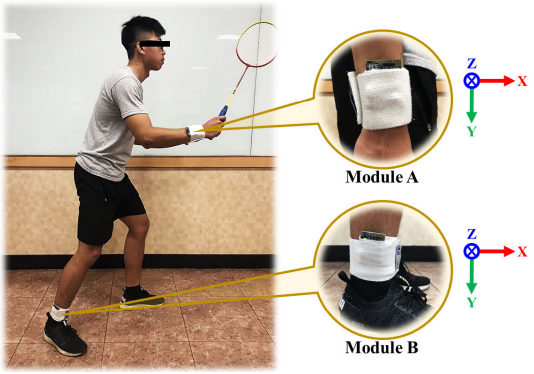


FIGURE 1. Wearable sport activity classification system mounted on an athlete's dominant wrist and ankle.

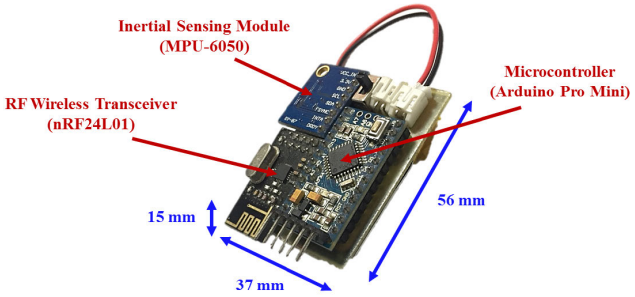


FIGURE 2. The proposed wearable inertial sensing module.

The experimental results and discussions are presented in Section IV. Finally, the conclusion is given in Section V.

II. WEARABLE SPORT ACTIVITY CLASSIFICATION SYSTEM

The proposed wearable sport activity classification system shown in Fig. 1 was designed in order to collect motion signals of sport activities, which consists of two wearable inertial sensing modules worn on the subjects' dominant wrist and ankle. The wearable inertial sensing module as shown in Fig. 2 was measured 56 mm × 37 mm × 15 mm with 16 grams. The wearable inertial sensing module is composed of a microcontroller (Arduino Pro Mini), a six-axis inertial sensor module (MPU-6050), an RF wireless transmission module (nRF24L01), and a power supply circuit. The schematic diagram of the wearable sport activity classification system is shown in Fig. 3. In this paper, we utilize an Arduino Pro Mini device as the microcontroller embedded in the wearable inertial sensing module, which is based on an ATmega328 core operating at a frequency of 16 MHz and 32 Kbytes of flash memory. The microcontroller collected the sport motion signals measured from the six-axis inertial sensor module (MPU-6050) through an I^2C interface and then transmitted the signals to a personal computer (PC) via the RF wireless transmission module (nRF24L01) through an SPI interface. The six-axis inertial sensor module, including a triaxial accelerometer, a triaxial gyroscope, and 16 bit analog to digital converters (ADCs), is applied to simultaneously

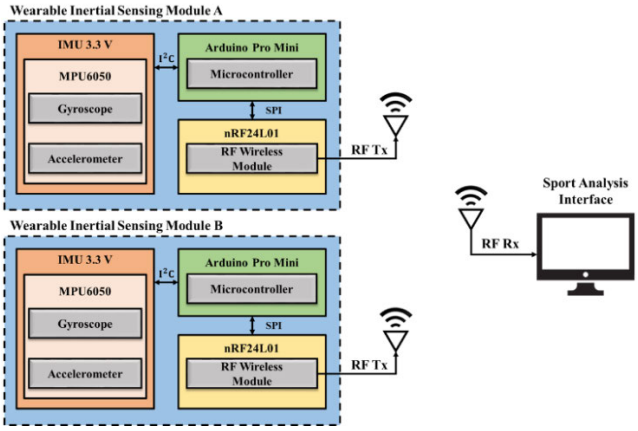


FIGURE 3. Schematic diagram of the wearable sport activity classification system.

collect the accelerations and angular velocities of the sport activities in a three-dimensional space and output the digital sport motion signals. The triaxial accelerometer is used to measure the gravitational and motion accelerations of hand and foot motions during executing sport activities and possesses a user selectable full scale of ±2, ±4, ±8, and ±16 g. The gyroscope can detect the angular velocities of hand and foot motions generated from sport activity movement and has a full scale of ±250, ±500, ±1000, and ±2000 °/s. In this paper, the sensitivity and range of the accelerometer were set as 2048 LSB/g and ±16 g, while which of the gyroscope were configured as 16.4 LSB/°/s and ±2000 °/s. The sampling rate of the accelerometer and gyroscope are both set as 100 Hz. The power consumption of the wearable inertial sensing module is provided by the power supply circuit consisting of a Li-ion battery, a Li-ion battery charging module, and regulators. The wearable inertial sensing module is powered by a 3.7 V polymer Li-ion battery with a nominal capacity of 450 mAh.

III. DEEP LEARNING-BASED SPORT ACTIVITY CLASSIFICATION ALGORITHM

In this paper, a deep learning-based sport activity classification algorithm has been developed for recognizing sport activities by using the accelerations and angular velocities measured by the accelerometer and gyroscope. The deep learning-based sport activity classification algorithm is composed of the following procedures: 1) sport motion signal collection, 2) signal preprocessing, 3) sport motion segmentation, 4) signal normalization, 5) spectrogram generation, 6) image mergence/resizing, and 7) CNN-based classification. First, the wearable inertial sensing modules worn on subjects' wrist and ankle collect the sport motion signals and transmit them to the PC-based sport analysis interface through the RF wireless transmission module. Second, in the signal preprocessing procedure, the processes of the calibration, lowpass filtering, and highpass filtering are carried out to reduce the sensitivity and offset errors from the measurements generated by the accelerometer and gyroscope,

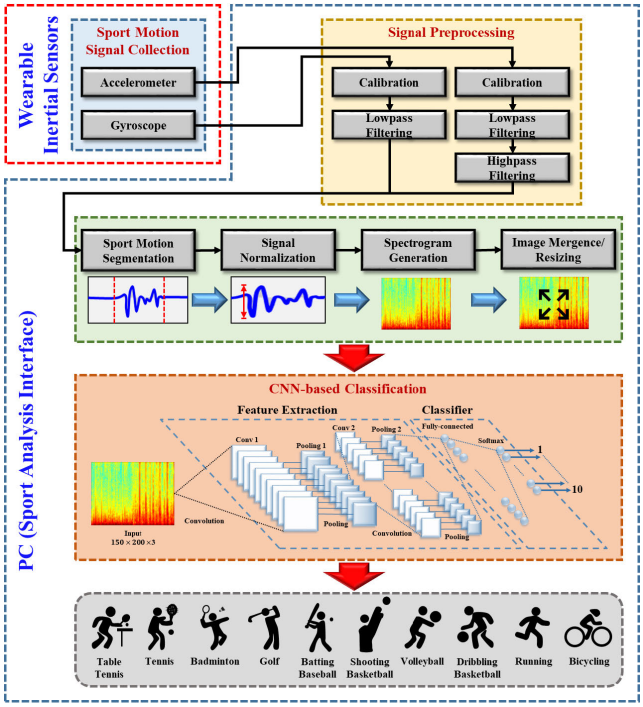


FIGURE 4. The proposed deep learning-based sport activity classification algorithm.

remove the users' unconscious trembles, and eliminate the gravitational acceleration, respectively. Third, an adaptive magnitude threshold method is developed to obtain sport motion intervals, which is based on the magnitude thresholds of the signal vector magnitudes of the filtered accelerations and angular velocities. Fourth, the size and amplitude of the sport motion signals within the sport motion intervals are normalized. Subsequently, the short-term Fourier transform (STFT) analysis is applied to the normalized accelerations and angular velocities for generating the spectrograms of the sport motion signals. Later, the spectrograms of the normalized accelerations and angular velocities are merged together to produce a new spectrogram, and then we perform spectrogram resize where set to 150 × 200 pixels. Finally, a CNN can be used for extracting the characteristic features from the resized spectrograms and classifying the ten types of sport activities, which comprises two convolutional layers, two pooling layers, a fully-connected layer, and a softmax layer. Fig. 4 shows the block diagram of the proposed deep learning-based sport activity classification algorithm. The details of the proposed deep learning-based sport activity classification algorithm are now described below.

A. SIGNAL PREPROCESSING

Both the accelerometer and gyroscope measurements always contain contamination of sensitivity and offset errors, environmental effects, and users' unconscious trembles in real-world applications. Therefore, the signal preprocessing composed of calibration and lowpass filter is essential in order to obtain the accurate measurements of the inertial sensors before the applications [34], [35]. On the other hand, the gravitational component of the measured accelerations is needed to be removed for acquiring the motion accelerations generated from the sport activities.

In this paper, the measurements of the triaxial accelerometer and gyroscope are calibrated firstly to reduce sensitivity and offset errors from the inertial sensors' raw measurements. For the accelerometer calibration, we first place the triaxial accelerometer on a leveled surface. Subsequently, each axis of the triaxial accelerometer are placed alternately upward and downward to align with the Earth's gravity which is only measured by the accelerometer when the wearable inertial sensing module is stationary. After then, the scale factors and offsets for all axes of the accelerometer can be obtained, which can be utilized to calibrate the raw measurements of the accelerometer as equation (1). For the gyroscope calibration, the sensitivity values for all axes represented in the datasheet can be used to be the scale factors of the gyroscope. Then, the mean of the angular velocities collected by keeping the wearable inertial sensing module stationary at the beginning are used to be the offsets for all axes of the triaxial gyroscope. Finally, we can calibrate the raw measurements of the gyroscope by using the scale factors and offsets as equation (1).

$$\mathbf{M}_c = \begin{bmatrix} SF_x & 0 & 0 \\ 0 & SF_y & 0 \\ 0 & 0 & SF_z \end{bmatrix} \times \mathbf{M}_r + \begin{bmatrix} O_x \\ O_y \\ O_z \end{bmatrix}, \quad (1)$$

where $\mathbf{M}_c$ is the calibrated accelerations ($\mathbf{a}_c = [a_{cx}\ a_{cy}\ a_{cz}]^T$) or angular velocities ($\boldsymbol{\omega}_c = [\omega_{cx}\ \omega_{cy}\ a\omega_{cz}]^T$). $\mathbf{M}_r$ is the raw accelerations ($\mathbf{a}_r = [a_{rx}\ a_{ry}\ a_{rz}]^T$) or angular velocities ($\boldsymbol{\omega}_r = [\omega_{rx}\ \omega_{ry}\ \omega_{rz}]^T$). SF_x, SF_y, and SF_z are the X-, Y-, and Z-axis scale factor of the triaxial accelerometer or gyroscope. O_x, O_y, and O_z are the X-, Y-, and Z-axis offset of the triaxial accelerometer or gyroscope. The more detailed information for the calibration procedures for the accelerometer and gyroscope can be found in [7] and [35].

Subsequently, the high frequency noise and users' unconscious trembles of the calibrated accelerations or angular velocities are further reduced by using a digital lowpass filter. In this paper, according to our empirical tests, we used a 15-point and a 5-point moving average filter for the sport motion signals generated from the hand and foot movements, respectively. Finally, the gravitational acceleration contained in the lowpass-filtered accelerations should be eliminated by using a digital highpass filter [36]. In this paper, the lowpass-filtered acceleration is passed through a three-order highpass elliptic filter with cut-off frequency of 0.005 Hz to remove the gravitational acceleration component for obtaining the motion accelerations generated by sport activities.

B. SPORT MOTION SEGMENTATION

The adaptive magnitude threshold method proposed in [4], consisting of calculation of signal vector magnitudes, calculation of adaptive magnitudes of signal vector magnitudes, finding start points of sport activity motion, and finding end points of sport activity motion, is utilized to automatically

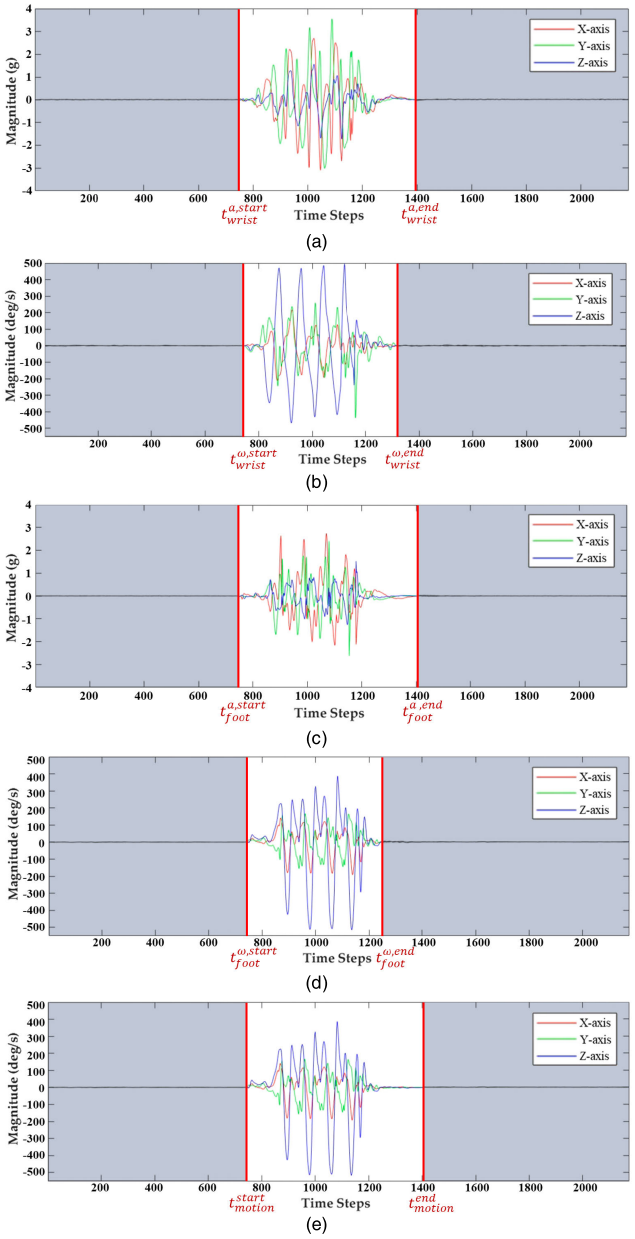


FIGURE 5. The partition of the sport motion interval by the proposed sport motion segmentation method for periodic sports. (a) Detection of the $t_{wrist}^{a,start}$ and $t_{wrist}^{a,end}$ for the hand acceleration. (b) Detection of the $t_{wrist}^{\omega,start}$ and $t_{wrist}^{\omega,end}$ for the hand angular velocity. (c) Detection of the $t_{foot}^{a,start}$ and $t_{foot}^{a,end}$ for the foot acceleration. (d) Detection of the $t_{foot}^{\omega,start}$ and $t_{foot}^{\omega,end}$ for the foot angular velocity. (e) Partition of the sport motion interval of the foot angular velocity into sport motion and static phases based on t_{motion}^{start} and t_{motion}^{end} simultaneously.

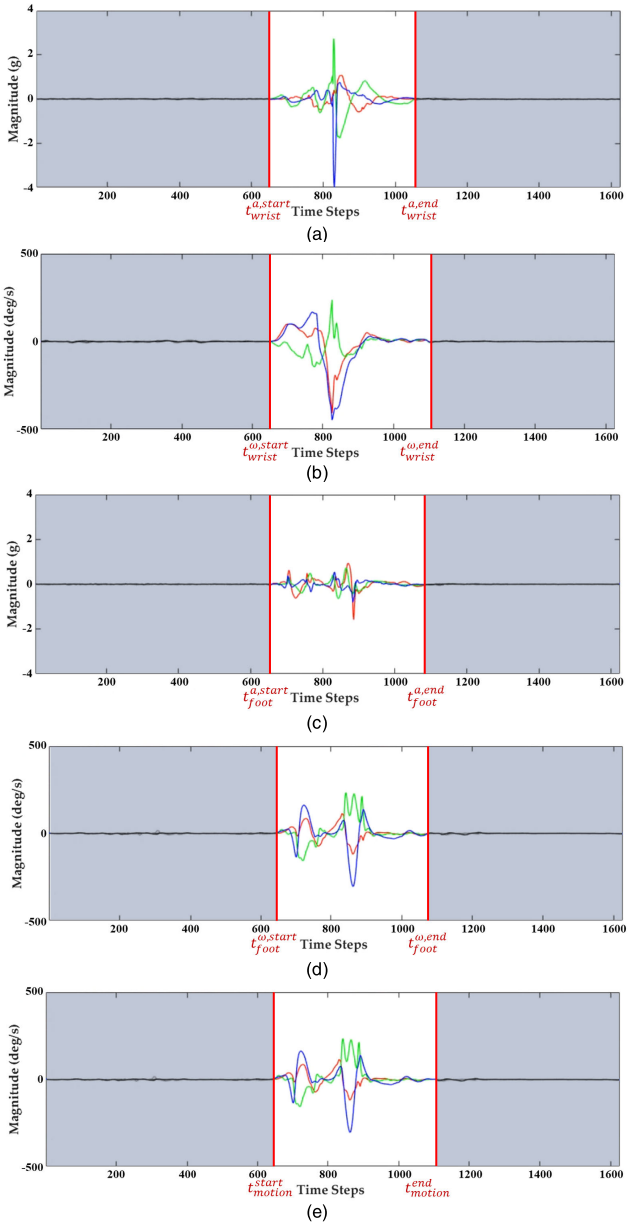


FIGURE 6. The partition of the sport motion interval by the proposed sport motion segmentation method for aperiodic sports. (a) Detection of the $t_{wrist}^{a,start}$ and $t_{wrist}^{a,end}$ for the hand acceleration. (b) Detection of the $t_{wrist}^{\omega,start}$ and $t_{wrist}^{\omega,end}$ for the hand angular velocity. (c) Detection of the $t_{foot}^{a,start}$ and $t_{foot}^{a,end}$ for the foot acceleration. (d) Detection of the $t_{foot}^{\omega,start}$ and $t_{foot}^{\omega,end}$ for the foot angular velocity. (e) Partition of the sport motion interval of the foot angular velocity into sport motion and static phases based on t_{motion}^{start} and t_{motion}^{end} simultaneously.

segment sport motion of each sport activity from the filtered accelerations and angular velocities generated from hand and foot motions. A detailed sport motion segmentation procedure can be found in [4]. The filtered accelerations and angular velocities measured from the wearable inertial sensing modules mounted on users' wrist and ankle for periodic (running) and aperiodic (tennis) sports are showed in Figs. 5 and 6. Figs. 5 and 6 are used to illustrate how to segment the sport motion interval for running and tennis by using the proposed sport motion segmentation method. The region with no color and the region with a grey color represent the sport motion and static intervals, respectively. From Figs. 5 and 6, we can found that the proposed sport motion segmentation method is reliable for all periodic and aperiodic sports.

C. SIGNAL NORMALIZATION

Once the sport motion interval has been segmented through the sport motion segmentation procedure, the size and amplitude of the sport motion signals collected by different users are inconsistent. To be able to eliminate the effect of human differences, each sport motion signal within the sport motion interval has been normalized into the size of the signal with the longest length via interpolation. Subsequently, we utilize the Z-score method to normalize the sport motion signals for reducing the single magnitude differences between the users. The Z-score normalization equation is shown as follows:

$$SM_n(k) = \frac{SM(k) - SM_{mean}}{SM_{std}}, \tag{2}$$

where $SM(k)$ and $SM_n(k)$ are the original and normalized sport motion signals within the sport motion interval, respectively, and k is the time steps in the sport motion interval. SM_{mean} and SM_{std} are the mean and standard deviation of the sport motion signals within the sport motion interval, respectively.

D. SPECTROGRAM GENERATION

After the procedure of signal normalization, the short-time Fourier transform (STFT) is applied to generate the spectrograms of the normalized sport motion signals. The spectrogram representation is a suitable domain for time-series data when a deep learning-based classifier is applied, which is indispensable to obtain interpretable features that can represent the intensity differences among nearest measurement data points [37]. The spectrogram of the normalized sport motion signal is the magnitude squared of the STFT of the normalized sport motion signal, which is a time-frequency signal representation. The computational procedure of the spectrogram is firstly to use a window function to divide a longer time signal into short segments with equal length and then compute the Fourier transform separately of each segment. The STFT of the normalized sport motion signal is defined as follows:

$$\begin{aligned} STFT\{x[n]\}(m,\omega) &= X(m,\omega) \\ &= \sum\nolimits_{n=-\infty}^{\infty} x[n]\omega[n-m]e^{-j\omega n}, \end{aligned} \tag{3}$$

where $x[n]$ is the normalized sport motion signals (SM_n), $\omega[n]$ is the window function, and m is the center of the window function. In this paper, the STFT is computed using a 64-point fast Fourier transform with a Hamming window overlapping with 50%. Subsequently, the spectrogram is the magnitude squared of the STFT of the normalized sport motion signals, which is defined as follows:

$$spectrogram\{x[n]\}(m,\omega) = |X(m,\omega)|^2. \tag{4}$$

The examples of the spectrograms across different sport activities are shown in Fig. 7.

E. IMAGE MERGENCE/RESIZING

Once the spectrograms of the X-, Y-, and Z-axis normalized sport motion signals measured by the accelerometers and gyroscopes mount on user's wrist and ankle are obtained via the spectrogram generation procedure, an image mergence method is utilized to merge the spectrograms of normalized sport motion signals generated from all axes of all inertial sensors. In the procedure of the spectrogram mergence, the spectrograms obtained from the X, Y, and Z axes are grouped column wise while the spectrograms computed on different sensors are grouped together row wise. After then, we magnify the merged spectrograms via nearest neighbor interpolation, which employs the value of the nearest point and ignores the values of other neighboring points, yielding a piecewise-constant interpolation [38]. In this study, each merged spectrogram is resized to 150 × 200 pixels for providing the inherent important features for the CNN classifier.

F. CNN-BASED CLASSIFICATION

Once the merged and resized spectrogram is obtained, each spectrogram image can be considered as the input data of the CNN classifier. The CNN classifier can divide the sport activities into 1) table tennis (backhand block), 2) tennis (forehand flat), 3) badminton (forehand clear), 4) golf (swing), 5) batting baseball (batting swing), 6) shooting basketball (set shot), 7) volleyball (underhand pass), 8) dribbling basketball, 9) running, and 10) bicycling, respectively. The CNN is utilized in this study for recognizing the ten sport activities, which is typically composed of one or more convolutional and pooling layers followed by several fully-connected layers. The CNN has a number of convolutional layers that can act as a feature extractor in order to obtain the high-level features from input data. The overall structure of the CNN is described below:

1) CONVOLUTIONAL LAYER

The convolutional layers use sliding filters to perform a convolution operation with the input spectrogram image. Subsequently, the values computed by the convolution operation are used to form the outputs of the convolutional layers, whose aspect ratio is roughly the same as that of the inputs of the convolutional layers. The outputs of the convolutional layers are referred to feature maps [39] and shown as follows:

$$c_{i,j}^{l,k} = f\left(b_k^l + \sum\nolimits_{m=1}^{M}\sum\nolimits_{n=1}^{N} w_{m,n}^{l,k} x_{i+m-1,j+n-1}^{l-1,k}\right), \tag{5}$$

where l is the layer index, M and N are the length and width of the kernel/filter size, b_k^l is the bias for the kth feature map of the lth layer, $w_{m,n}^{l,k}$ is the weight between the input $x_{i+m-1,j+n-1}^{l-1,k}$ and the kth feature map of the lth layer, and f is the activation function where the ReLU is adopted as the activation function in this paper, which can be expressed as follows:

$$f(x) = \max(0, x). \tag{6}$$

2) POOLING LAYER

the pooling layer usually follows the convolutional layer to perform a summary statistic of nearby outputs, which is

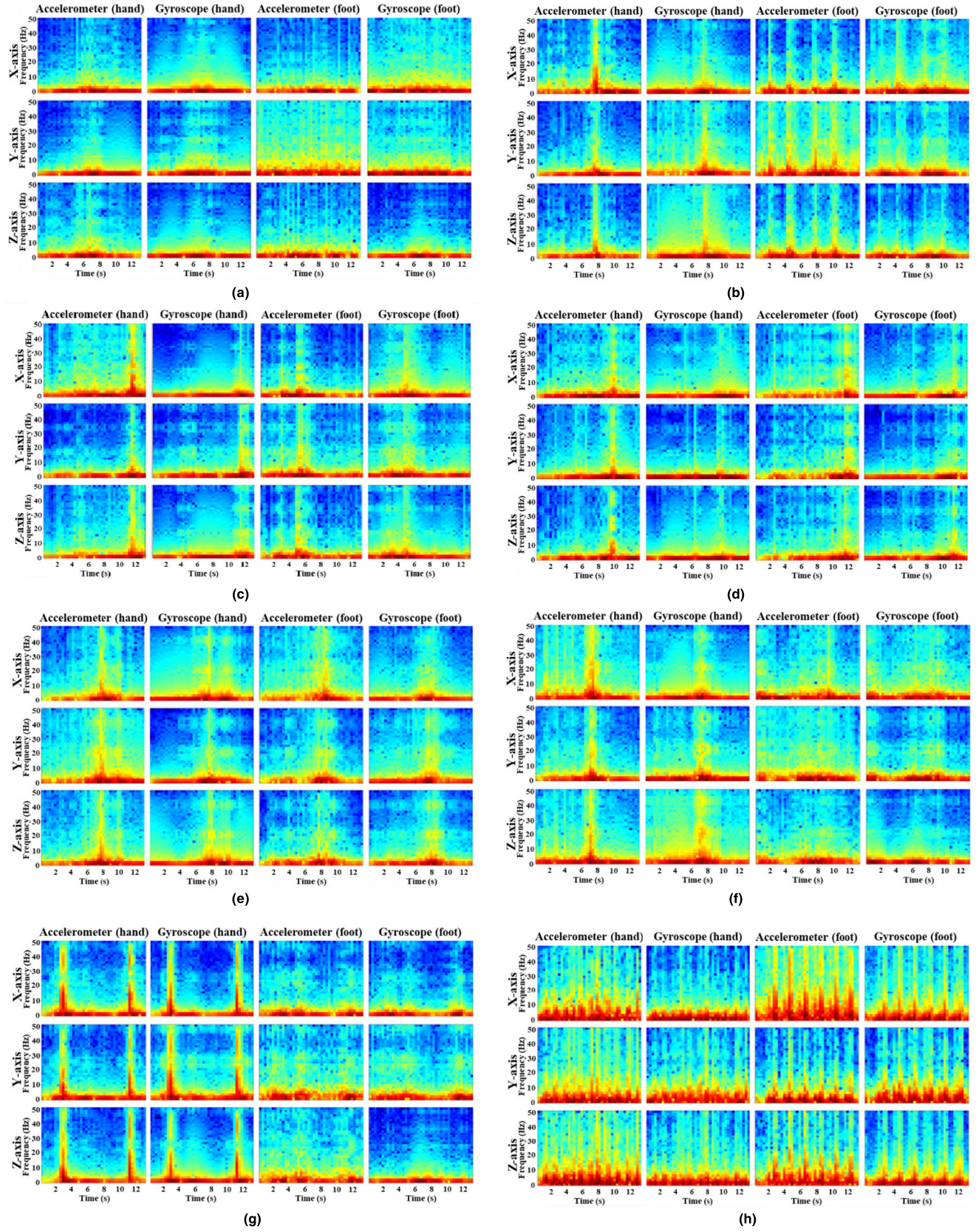


FIGURE 7. Examples of spectrogram extracted from the ten sport activities. (a) table tennis, (b) tennis, (c) badminton, (d) golf, (e) batting baseball, (f) shooting basketball, (g) volleyball, (h) dribbling basketball, (i) running, and (j) bicycling.

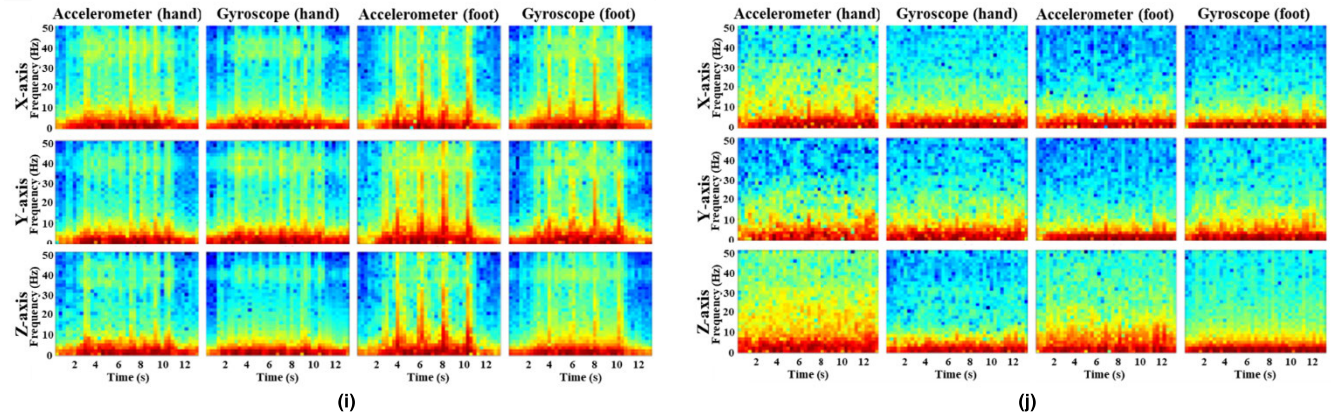


FIGURE 7. *(Continued.)* **Examples of spectrogram extracted from the ten sport activities. (a) table tennis, (b) tennis, (c) badminton, (d) golf, (e) batting baseball, (f) shooting basketball, (g) volleyball, (h) dribbling basketball, (i) running, and (j) bicycling.**

derived from $c_{i,j}^{l,k}$. The pooling operation used in this paper is max-pooling, which outputs the maximum value among nearby feature maps, and can be calculated as follows:

$$p_{i,j}^{l,k} = \max_{q \in Q} (c_{i \times R+q, j \times R+q}^{l,k}), \tag{7}$$

where $Q = 2 \times 2$ is the pooling size and $R = 2$ is pooling stride in this paper.

3) FULLY-CONNECTED LAYER

Inherent features extracted from the stacked convolutional and pooling layers are flattened into a one-dimensional feature vector as the input for the fully-connected layer:

$$o_h^l = f(\sum\nolimits_g (w_{hg}^l p_g^{l-1}) + b_h^l), \tag{8}$$

where w_{hg}^l is the weight between the gth node on layer $l-1$ and hth node on layer l, b_h^l is the bias for the hth node on the lth layer, and f is the relu activation function.

4) SOFTMAX LAYER

the output of the last layer, softmax layer, computes probability distribution over the predicted sport activity classes.

$$P(c \mid o) = \underset{c \in C}{\operatorname{argmax}} \frac{\exp(o_h^L)}{\sum_{m=1}^{N_C} \exp(o_m^L)}, \tag{9}$$

where c is the sport activity class, N_C is the total number of sport activity classes, and L is the last layer index. in the training phase, we utilize the stochastic gradient descend (SGD) algorithm to update the network parameters for minimizing a categorical cross-entropy loss function [28].

In this paper, the structure of CNN classifier is determined by a trial-and-error process. The proposed CNN classifier shown in Fig. 4 contains two convolutional layers, two pooling layers, one fully-connected layer, and one softmax layer. Each convolutional layer followed by an ReLU activation function and a 2×2 max-pooling layer. The first convolutional layer uses the $150 \times 200 \times 3$ image and applies ten 1×10 filters. Subsequently, the volume of the original image becomes $75 \times 95 \times 10$ through an ReLU and a 2×2 max-pooling layer. The second convolutional layer has ten 1×10 filters and is followed by an ReLU and a 2×2 max-pooling layer, resulting in a $37 \times 43 \times 10$ image volume. Once all convolution and max-pooling are performed, the $37 \times 43 \times 10$ image volume is flattened into a 15910×1 feature vector in the fully-connected layer. The fully-connected layer reduces the size of the feature vector from 15910 to 10, and then the softmax layer takes the 10×1 reduced feature vector and outputs the final classification result. In this paper, the output of the CNN classifier is represented as the label of the ten types of sport activities (i.e., table tennis (S1), tennis (S2), badminton (S3), golf (S4), batting baseball (S5), shooting basketball (S6), volleyball (S7), dribbling basketball (S8), running (S9), and bicycling (S10) are labeled as '1', '2', '3', '4', '5', '6', '7', '8', '9', and '10') for the sport activity classification task.

TABLE 1. Participant characteristics for sport activity classification experiment.

Parameters	N = 10
Men/Women (n)	6/4
Age (years)	21.4 ± 0.8
Height (cm)	168.7 ± 7.2
Weight (kg)	60.1 ± 12.7
BMI (kg/m^2)	21.0 ± 3.2

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Our experiments were performed on a PC running Microsoft Windows 10 operating system with an Intel®Core Processor i7-7700 and 16-GB RAM that drive a NVIDIA GeForce GTX 1080 Ti GPU. The video observation was regarded as the criterion measure which was used to validate the classification results classified by the proposed deep learning-based sport activity classification algorithm. The measures of accuracy (*Acc*), specificity (*Sp*), sensitivity (*Se*), and correct classification rate (*CCR*) were utilized to evaluate the performances of the proposed deep learning-based sport activity classification

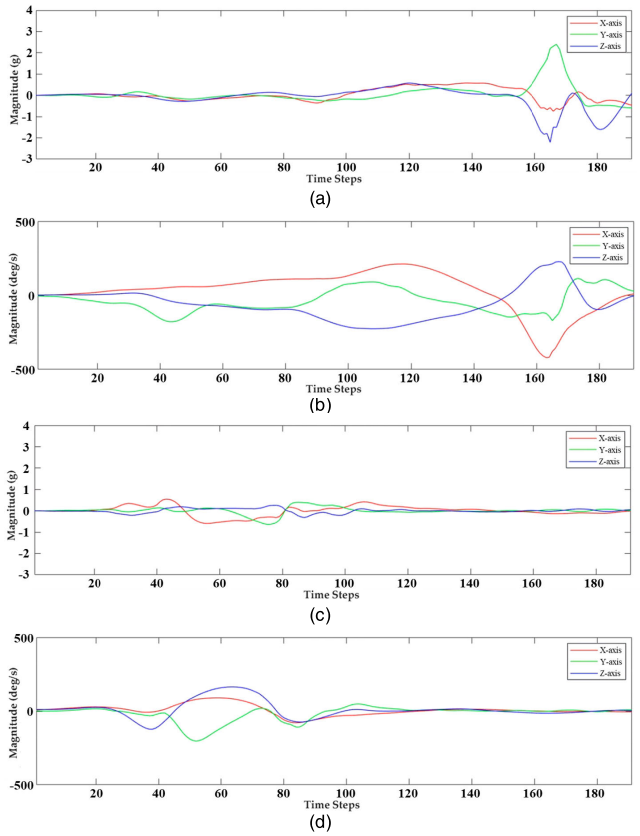


FIGURE 8. Motion signals generated from the badminton activity. (a) Hand accelerations. (b) Hand angular velocities. (c) Foot accelerations. (d) Foot angular velocities.

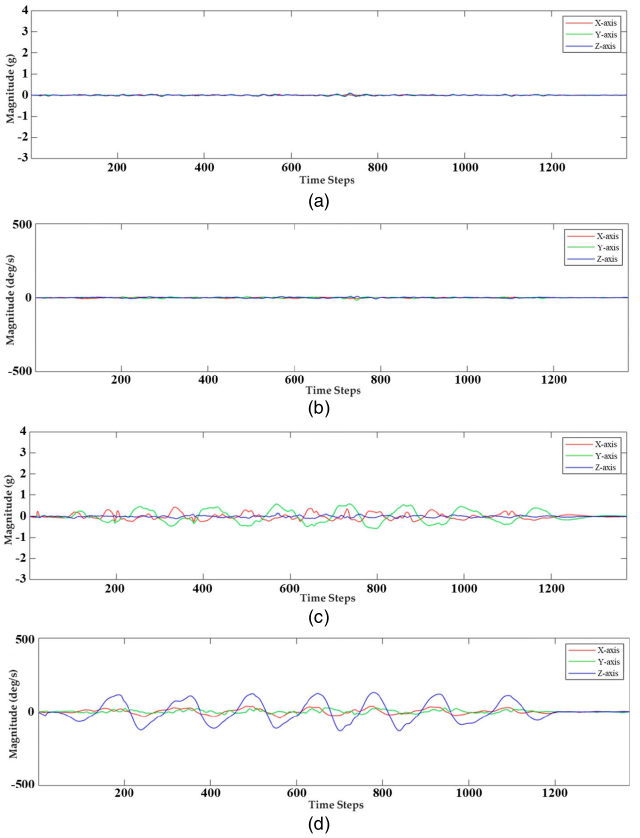


FIGURE 9. Motion signals generated from the bicycling activity. (a) Hand accelerations. (b) Hand angular velocities. (c) Foot accelerations. (d) Foot angular velocities.

algorithm, and defined as follows:

$$Acc(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100, \quad (10)$$

$$S_p(\%) = \frac{TN}{TN + FP} \times 100, \quad (11)$$

$$S_e(\%) = \frac{TP}{TP + FN} \times 100, \quad (12)$$

$$CCR(\%) = \frac{\sum_{i=1}^{C} TP_i}{TP + TN + FP + FN} \times 100, \quad (13)$$

where TP, TN, FP, and FN are true positive, true negative, false positive, and false negative, respectively. The overall classification performance is evaluated by CCR, where C is the number of classes. In addition, we evaluated the classification performances of the deep learning-based sport activity classification algorithm by using the 2-fold, 5-fold, 10-fold, leave-one-out (LOO), and leave-one-subject-out (LOSO) cross-validation strategies.

A. PARTICIPANTS AND EXPERIMENTAL SETUP

In this paper, ten subjects (six males and four females; age = 21.4 ± 0.8 years; height = 168.7 ± 7.2 cm; mass = 60.1 ± 12.7 kg) without any self-reported injuries or musculoskeletal disorders were recruited to collect sport motion signals for the sport activity classification experiment. All participants were instructed to perform 10 types of sport activities with each activity executed 10 times. In addition, the participants were asked to execute each sport activity lasting about 15 s, and allowed to rest if needed. For all aperiodic sports (S1-S7), the sport movement for each sport is conducted once only in 15s and then back to relax status. On the other hand, the sport movement for each periodic sport (S8-S10) is executed repeatedly and continuously lasting about 15s. Therefore, a total number of 1000 (= 10 × 10 × 10) samples were used to validate the classification performances of the proposed wearable sport activity classification system and its associated deep learning-based sport activity classification algorithm. Table 1 shows the demographics of the participants for the sport activity classification experiment. The experimental protocol used in this paper was approved by Institutional Review Board (IRB) of the National Cheng Kung University Hospital. All subjects were required to provide written informed consent before study participation.

B. SPORT ACTIVITY CLASSIFICATION

In this experiment, each participant was invited to mount the wearable inertial sensor modules on their wrist and ankle and execute sport activities in a laboratory environment. Figs. 8 and 9 show the motion signal generated from the

TABLE 2. Classification performance comparisons of the proposed deep learning-based sport activity classification algorithm by cross-validation in sport activity recognition task.

K-fold \ Performance	Acc (%)	Sp (%)	Se (%)	CCR (%)
2-fold	99.48	99.71	97.40	97.40
5-fold	99.76	99.87	98.80	98.80
10-fold	99.86	99.92	99.30	99.30
LOO	99.80	99.89	99.00	99.00
LOSO	99.64	99.80	98.20	98.20

TABLE 3. Classification performance comparisons of CCRs for all 10 sports.

K-fold \ Sport	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Average CCR
2-fold	99%	94%	96%	93%	94%	99%	99%	100%	100%	100%	97.40%
5-fold	99%	97%	98%	98%	97%	99%	100%	100%	100%	100%	98.80%
10-fold	99%	98%	100%	99%	98%	100%	99%	100%	100%	100%	99.30%
LOO	99%	100%	98%	96%	97%	100%	100%	100%	100%	100%	99.00%
LOSO	99%	96%	99%	96%	95%	98%	100%	99%	100%	100%	98.20%

TABLE 4. Confusion matrix for sport recognition using the proposed deep learning-based sport activity classification algorithm verified by 2-fold cross-validation.

Activity \ Classified	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	99	0	0	0	0	1	0	0	0	0
S2	0	94	1	2	2	1	0	0	0	0
S3	0	4	96	0	0	0	0	0	0	0
S4	2	2	1	93	2	0	0	0	0	0
S5	1	2	0	2	94	1	0	0	0	0
S6	0	0	0	0	1	99	0	0	0	0
S7	0	0	0	0	0	1	99	0	0	0
S8	0	0	0	0	0	0	0	100	0	0
S9	0	0	0	0	0	0	0	0	100	0
S10	0	0	0	0	0	0	0	0	0	100

TABLE 5. Confusion matrix for sport recognition using the proposed deep learning-based sport activity classification algorithm verified by 5-fold cross-validation.

Activity \ Classified	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	99	0	0	0	0	1	0	0	0	0
S2	0	97	1	1	0	1	0	0	0	0
S3	0	1	98	1	0	0	0	0	0	0
S4	0	0	1	98	1	0	0	0	0	0
S5	0	0	0	1	97	1	1	0	0	0
S6	0	0	0	0	1	99	0	0	0	0
S7	0	0	0	0	0	0	100	0	0	0
S8	0	0	0	0	0	0	0	100	0	0
S9	0	0	0	0	0	0	0	0	100	0
S10	0	0	0	0	0	0	0	0	0	100

aperiodic (badminton) and periodic (bicycling) sport activities, respectively. The overall accuracy of the proposed deep learning scheme was achieved 99.48%, 99.76%, 99.86%, 99.80%, and 99.64% by the 2-fold, 5-fold, 10-fold, LOO, and LOSO cross-validation, respectively. The overall CCRs of the proposed deep learning scheme were achieved 97.40%, 98.80%, 99.30%, 99.00%, and 98.20% by the 2-fold, 5-fold, 10-fold, LOO, and LOSO cross-validation, respectively. The classification performance comparisons of the proposed deep learning scheme with CNN classifier by the 2-fold, 5-fold, 10-fold, LOO, and LOSO cross-validation strategies are summarized in Table 2. Obviously, these results demonstrate that the proposed deep learning scheme can obtain good performance for classifying the sport activities. In addition, the performance comparisons of CCRs for all 10 sports are shown in Table 3, which shows that golf (S4) and batting baseball (S5) have worse classification rates than other sports. In addition, we can find that the CCRs of the periodic sports (S8-S10) are better than that of the aperiodic sports (S1-S7). The confusion matrices of the sport activity classification using the proposed deep learning scheme verified by the 2-fold, 5-fold, 10-fold, LOO, and LOSO cross-validation

TABLE 6. Confusion matrix for sport recognition using the proposed deep learning-based sport activity classification algorithm verified by 10-fold cross-validation.

Activity \ Classified	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	99	0	0	0	1	0	0	0	0	0
S2	0	98	1	1	0	0	0	0	0	0
S3	0	0	100	0	0	0	0	0	0	0
S4	0	1	0	99	0	0	0	0	0	0
S5	0	0	0	1	98	0	1	0	0	0
S6	0	0	0	0	0	100	0	0	0	0
S7	0	0	0	0	0	1	99	0	0	0
S8	0	0	0	0	0	0	0	100	0	0
S9	0	0	0	0	0	0	0	0	100	0
S10	0	0	0	0	0	0	0	0	0	100

TABLE 7. Confusion matrix for sport recognition using the proposed deep learning-based sport activity classification algorithm verified by LOO cross-validation.

Activity \ Classified	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	99	0	0	0	0	1	0	0	0	0
S2	0	100	0	0	0	0	0	0	0	0
S3	0	2	98	0	0	0	0	0	0	0
S4	0	2	0	96	2	0	0	0	0	0
S5	1	0	0	1	97	0	1	0	0	0
S6	0	0	0	0	0	100	0	0	0	0
S7	0	0	0	0	0	0	100	0	0	0
S8	0	0	0	0	0	0	0	100	0	0
S9	0	0	0	0	0	0	0	0	100	0
S10	0	0	0	0	0	0	0	0	0	100

TABLE 8. Confusion matrix for sport recognition using the proposed deep learning-based sport activity classification algorithm verified by LOSO cross-validation.

Activity \ Classified	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
S1	99	0	0	0	0	1	0	0	0	0
S2	0	96	1	2	1	0	0	0	0	0
S3	0	0	99	0	0	1	0	0	0	0
S4	0	1	1	96	1	0	0	1	0	0
S5	1	1	0	2	95	0	1	0	0	0
S6	0	0	2	0	0	98	0	0	0	0
S7	0	0	0	0	0	0	100	0	0	0
S8	0	0	0	0	0	0	0	99	0	1
S9	0	0	0	0	0	0	0	0	100	0
S10	0	0	0	0	0	0	0	0	0	100

TABLE 9. Classification performance comparisons of the proposed scheme with some existing methods for sport activity classification.

Reference	Sensor	Placement	No. of Sport Activity	No. of Feature	Classifier	Acc (%)	CCR (%)
Proposed method	Accelerometer, Gyroscope	Wrist, Ankle	10	--	CNN	99.86 (10-fold)	99.30 (10-fold)
[25]	Accelerometer, GPS	Wrist, Hip	9	7	Custom decision tree + ANN	89.00 (12-fold)	--
[24]	Accelerometer	Wrist	8	11	Logistic regression/ANN	85.00	--
				--	Dynamic time warping (DTW)	80.00	--
[40]	Camera	--	10	--	CNN	95.06	--

strategies are shown in Tables 4-8. In terms of the runtime of the training processes, the proposed deep learning-based sport activity classification algorithm spent 43.6 s, 255 s, 516 s, 50896.5 s, and 508 s by the 2-fold, 5-fold, 10-fold, LOO, and LOSO cross-validation strategies, respectively. Furthermore, the runtime of the test process was 0.0024 s. From Table 4 to Table 8, we can find that the most misclassification occurred at golf (S4) and batting baseball (S5). The reason is that these two sport activities have identical characteristic features which are aperiodic hand swing and foot-based motions. Obviously, from Figs. 10 and 11, the hand accelerations of these two sport activities are similar and

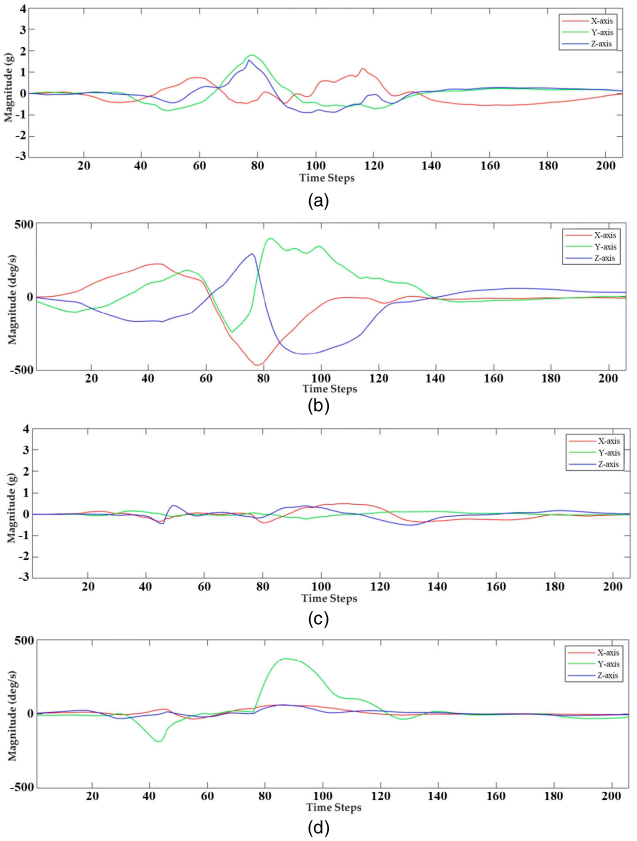


FIGURE 10. Motion signals generated from the golf activity. (a) Hand accelerations. (b) Hand angular velocities. (c) Foot accelerations. (d) Foot angular velocities.

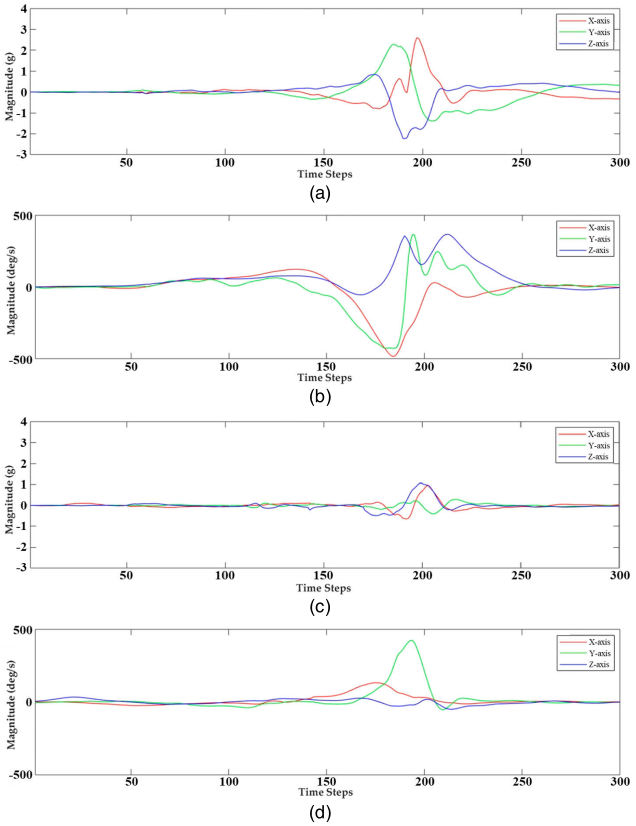


FIGURE 11. Motion signals generated from the batting baseball activity. (a) Hand accelerations. (b) Hand angular velocities. (c) Foot accelerations. (d) Foot angular velocities.

aperiodic, while the hand angular velocities of these two sport activities are also similar to each other.

C. DISCUSSION

In this section, we compared the classification performance of the proposed deep learning scheme for sport activity classification with those of other methods simultaneously utilizing different classification methods presented in the literature. We selected 3 existing sport activity classification methods for comparison with our proposed algorithm. The performance comparisons of our proposed algorithm and 3 existing methods for the sport activity classification task are summarized in Table 9. The classification results compare with the results reported in [24], [25], and [40] whose average sport activity classification accuracy are 85.00%, 89.00%, and 95.06%. The overall accuracy of the proposed CNN classifier was better than that of the logistic regression (LR) or ANN [24], the hybrid classifier combining the custom decision tree and ANN [25], and CNN [40] by more than 14.86%, 10.86%, and 4.8%, respectively. Therefore, the experimental results indicate that the proposed wearable sport activity classification system with the deep learning-based sport activity classification algorithm is effective to provide an accurate classification rate in the sport activity recognition task. Note that, the abovementioned studies utilized various classification methods to classify numerous sport activities with high accuracy; however, the performance of them would be affected by their individual datasets and number of sport activities. In addition, the limitations of this paper include: 1) the proposed method has been evaluated on exemplar samples of sport activities that were conducted in a laboratory environment, and 2) the proposed CNN-based sport activity classification algorithm is performed in offline. In addition, the size of the wearable inertial sensing modules will be minimized for improving wear comfortability. Furthermore, related to segmentation errors is the importance for recognizing sports involving multiple movements. In future research, we will develop a more accurate sport movement segmentation algorithm from sport movement signals in order to distinguish and recognize multiple movement patterns of sports.

V. CONCLUSION

In this paper, a wearable sport activity classification system and its associated deep learning-based sport activity classification algorithm are presented to accurately classify ten sport activities consisting of table tennis, tennis, badminton, golf, batting baseball, volleyball, shooting basketball, dribbling basketball, running, and bicycling. The proposed deep learning-based sport activity classification

algorithm composed of sport motion signal collection, signal preprocessing, sport motion segmentation, signal normalization, spectrogram generation, image mergence/resizing, and CNN-based classification is developed to classify ten types of sport activities. The inherent features are extracted from the spectrograms of the STFT of the sport motion signals using the CNN classifier which comprises two convolutional layers, two pooling layers, a fully-connected layer, and a softmax layer. The overall CCRs of 97.40%, 98.80%, 99.30%, 99.00%, and 98.20% by the 2-fold, 5-fold, 10-fold, LOO, and LOSO cross-validation strategies for ten sport activity classes, respectively, can be achieved by using the proposed deep learning scheme. Based on the abovementioned experimental results, the effectiveness of the proposed deep learning-based sport activity classification algorithm has been successfully validated. We believe that the proposed wearable sport activity classification system and its associated deep learning-based sport activity classification algorithm can be regarded as an effective method for inertial-sensing-based sport activity classification tasks.

REFERENCES

[1] R. J. Aughey and C. Falloon, "Real-time versus post-game GPS data in team sports," *J. Sci. Med. Sport*, vol. 13, no. 3, pp. 348–349, 2010.

[2] M. Cornacchia, K. Ozcan, Y. Zheng, and S. Velipasalar, "A survey on activity detection and classification using wearable sensors," *IEEE Sensors J.*, vol. 17, no. 2, pp. 386–403, Jan. 2017.

[3] E. E. Cust, A. J. Sweeting, K. Ball, and S. Robertson, "Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance," *J. Sports Sci.*, vol. 37, no. 5, pp. 568–600, 2019.

[4] Y.-L. Hsu, S.-C. Yang, H.-C. Chang, and H.-C. Lai, "Human daily and sport activity recognition using a wearable inertial sensor network," *IEEE Access*, vol. 6, pp. 31715–31728, 2018.

[5] G. Thomas, R. Gade, T. B. Moeslund, P. Carr, and A. Hilton, "Computer vision for sports: Current applications and research topics," *Comput. Vis. Image Understand.*, vol. 159, pp. 3–18, Jun. 2017.

[6] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim, "Bring it to the pitch: Combining video and movement data to enhance team sport analysis," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 13–22, Aug. 2018.

[7] H.-C. Chang, Y.-L. Hsu, S.-C. Yang, J.-C. Lin, and Z.-H. Wu, "A wearable inertial measurement system with complementary filter for gait analysis of patients with stroke or Parkinson's disease," *IEEE Access*, vol. 4, pp. 8442–8453, 2016.

[8] P. Musale, D. Baek, N. Werellagama, S. S. Woo, and B. J. Choi, "You walk, we authenticate: Lightweight seamless authentication based on gait in wearable IoT systems," *IEEE Access*, vol. 7, pp. 37883–37895, 2019.

[9] S. Qiu, L. Liu, Z. Wang, S. Li, H. Zhao, J. Wang, J. Li, and K. Tang, "Body sensor network-based gait quality assessment for clinical decision-support via multi-sensor fusion," *IEEE Access*, vol. 7, pp. 59884–59894, 2019.

[10] U. Rashid, N. Kumari, D. Taylor, T. David, and N. Signal, "Gait event anomaly detection and correction during a split-belt treadmill task," *IEEE Access*, vol. 7, pp. 68469–68478, 2019.

[11] Y. Tain, J. Zhang, L. Chen, Y. Geng, and X. Wang, "Single wearable accelerometer-based human activity recognition via kernel discriminant analysis and QPSO-KELM classifier," *IEEE Access*, vol. 7, pp. 109216–109227, 2019.

[12] Y. Zhao, J. Liang, X. Sha, J. Yu, H. Duan, G. Shi, and W. J. Li, "Estimation of pedestrian altitude inside a multi-story building using an integrated micro-IMU and barometer device," *IEEE Access*, vol. 7, pp. 84680–84689, 2019.

[13] V. Camomilla, E. Bergamini, S. Fantozzi, and G. Vannozzi, "Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review," *Sensors*, vol. 18, no. 3, Mar. 2018, Art. no. 873.

[14] M. Kos and I. Kramberger, "A wearable device and system for movement and biometric data acquisition for sports applications," *IEEE Access*, vol. 5, pp. 6411–6420, 2017.

[15] K. King, J. Hough, R. McGinnis, and N. C. Perkins, "A new technology for resolving the dynamics of a swinging bat," *Sports Eng.*, vol. 15, no. 1, pp. 41–52, 2012.

[16] H. Ghasemzadeh and R. Jafari, "Coordination analysis of human movements with body sensor networks: A signal processing model to evaluate baseball swings," *IEEE Sensors J.*, vol. 11, no. 3, pp. 603–610, Mar. 2011.

[17] Y. J. Kim, K. D. Kim, S. H. Kim, S. Lee, and H. S. Lee, "Golf swing analysis system with a dual band and motion analysis algorithm," *IEEE Trans. Consum. Electron.*, vol. 63, no. 3, pp. 309–317, Nov. 2017.

[18] F. Malawski and B. Kwolek, "Recognition of action dynamics in fencing using multimodal cues," *Image Vis. Comput.*, vol. 75, pp. 1–10, Jul. 2018.

[19] M. Mangiarotti, F. Ferrise, S. Graziosi, F. Tamburrino, and M. Bordegoni, "A wearable device to detect in real-time bimanual gestures of basketball players during training sessions," *J. Comput. Inf. Sci. Eng.*, vol. 19, no. 1, Mar. 2019, Art. no. 011004.

[20] Y. Wang, M. Chen, X. Wang, R. H. M. Chan, and W. J. Li, "IoT for next-generation racket sports training," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4558–4566, May 2018.

[21] Z. Wang, M. Guo, and C. Zhao, "Badminton stroke recognition based on body sensor networks," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 5, pp. 769–775, Oct. 2016.

[22] K. L. Wang and J. Xu, "A speed regression using acceleration data in a deep convolutional neural network," *IEEE Access*, vol. 7, pp. 9351–9356, 2019.

[23] J. Y. Xu, X. Nan, V. Ebkeb, Y. Wang, G. J. Pottie, and W. J. Kauser, "Integrated inertial sensors and mobile computing for real-time cycling performance guidance via pedaling profile classification," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 2, pp. 440–445, May 2015.

[24] J. Margarito, R. Helaoui, A. M. Bianchi, F. Sartor, and A. G. Bonomi, "User-independent recognition of sports activities from a single wrist-worn accelerometer: A template-matching-based approach," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 788–796, Apr. 2016.

[25] M. Ermers, J. Pärkkä, J. Mäntyjärvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 20–26, Jan. 2008.

[26] E. Mitchell, D. Monaghan, and N. E. O'Connor, "Classification of sporting activities using smartphone accelerometers," *Sensors*, vol. 13, no. 4, pp. 5317–5337, 2013.

[27] Y. Wang, Y. Zhao, R. H. M. Chan, and W. J. Li, "Volleyball skill assessment using a single wearable micro inertial measurement unit at wrist," *IEEE Access*, vol. 6, pp. 13758–13765, 2018.

[28] M. Gadaleta and M. Rossi, "IDNet: Smartphone-based gait recognition with convolutional neural networks," *Pattern Recognit.*, vol. 74, pp. 25–37, Feb. 2018.

[29] M. Munoz-Organero, "Human activity recognition based on single sensor square HV acceleration images and convolutional neural networks," *IEEE Sensors J.*, vol. 19, no. 4, pp. 1487–1498, Feb. 2019.

[30] S. Wan, Y. Liang, Y. Zhang, and M. Guizani, "Deep multi-layer perceptron classifier for behavior analysis to estimate Parkinson's disease severity using smartphones," *IEEE Access*, vol. 6, pp. 36825–36833, 2018.

[31] R. Zhu, Z. Xiao, Y. Li, M. Yang, Y. Tan, L. Zhou, S. Lin, and H. Wen, "Efficient human activity recognition solving the confusing activities via deep ensemble learning," *IEEE Access*, vol. 7, pp. 75490–75499, 2019.

[32] T. Kautz, B. H. Groh, J. Hannink, U. Jensen, H. Strubberg, and B. M. Eskofier, "Activity recognition in beach volleyball using a deep convolutional neural network," *Data Mining Knowl. Discovery*, vol. 31, no. 6, pp. 1678–1705, 2017.

[33] L. Jiao, R. Bie, H. Wu, Y. Wei, J. Ma, A. Umek, and A. Kos, "Golf swing classification with multiple deep convolutional neural networks," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 10, Oct. 2018, Art. no. 155014771 8802186.

[34] Y.-L. Hsu, J.-S. Wang, and C.-W. Chang, "A wearable inertial pedestrian navigation system with quaternion-based extended Kalman filter for pedestrian localization," *IEEE Sensors J.*, vol. 17, no. 10, pp. 3193–3206, May 2017.

[35] J.-S. Wang, Y.-L. Hsu, and J.-N. Liu, "An inertial-measurement-unit-based pen with a trajectory reconstruction algorithm and its applications," *IEEE Trans. Ind. Electron.*, vol. 57, no. 10, pp. 3508–3521, Oct. 2010.

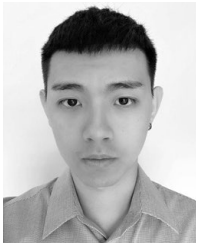
[36] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 156–167, Jan. 2006.
[37] D. Ravì, C. Wong, B. Lo, and G.-Z. Yang, "A deep learning approach to on-node sensor data analytics for mobile or wearable devices," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 56–64, Jan. 2017.
[38] T. Hongshen, R. Ni, Y. Zhao, and X. Li, "Median filtering detection of small-size image based on CNN," *J. Vis. Commun. Image Represent.*, vol. 51, pp. 162–168, Feb. 2018.
[39] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.
[40] X. Chen, J. Weng, W. Lu, J. Xu, and J. Weng, "Deep manifold learning combined with convolutional neural networks for action recognition," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 3938–3952, Sep. 2018.

YU-LIANG HSU (M'17) received the B.S. degree in automatic control engineering from Feng Chia University, Taichung, Taiwan, in 2004, and the M.S. and Ph.D. degrees in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 2007 and 2011, respectively.

He is currently an Associate Professor with the Department of Automatic Control Engineering, Feng Chia University, Taiwan. His research interests include computational intelligence, biomedical engineering, nonlinear system identification, and wearable intelligent technology.

HSING-CHENG CHANG received the B.S. and M.S. degrees in physics from Tamkang University, Taiwan, in 1978 and 1980, respectively, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Cincinnati, Cincinnati, OH, USA, in 1991 and 1994, respectively.

He is currently a Professor with the Department of Automatic Control Engineering, Feng Chia University, Taiwan. His research interests include microsensors and microactuators, circuit design, automation technology, and engineering education.



YUNG-JUNG CHIU received the B.S. and M.S. degrees in automatic control engineering from Feng Chia University, Taichung, Taiwan, in 2017 and 2019, respectively. His research interests include signal processing, wearable devices, and deep learning.