Human Daily Activity Recognition Performed Using Wearable Inertial Sensors Combined With Deep Learning Algorithms

CHIH-TA YEN, (Member, IEEE), JIA-XIAN LIAO, AND YI-KAI HUANG
Department of Electrical Engineering, National Formosa University, Huwei 632, Taiwan
Corresponding author: Chih-Ta Yen (chihtayen@gmail.com)

ABSTRACT This study proposed a wearable device capable of recognizing six human daily activities (walking, walking upstairs, walking downstairs, sitting, standing, and lying) through a deep learning algorithm. Existing wearable devices are mainly watches or wristbands, and almost none are to be worn on the waist. Wearable devices in the forms of watches and wristbands are unfriendly to patients who are critically ill, such as patients undergoing dialysis. Patients undergoing dialysis have artificial blood vessels on their arm, and they cannot perform intense exercise. For this type of users, general hand wearable devices cannot correctly identify wearers' activities. Therefore, we proposed a waist wearable device and these types of daily life activities to assess their exercise. The hardware of the wearable device consisted of an inertial sensor, which included a microcontroller, a three-axis accelerometer, and a three-axis gyroscope. The activity recognition algorithm of the software used motion signals acquisition, signal normalization, and a feature learning method. The feature learning method was based on a 1D convolutional neural network that automatically performed feature extraction and classification from raw data. One part of the experimental data was from the dataset of the University of California (UCI), and the other part was recorded by this study. To capture the data recorded, the wearable inertial sensing device was attached to the waists of 21 experimental participants who performed six common movements in a laboratorial environment, and the subsequent records were collected to verify the validity of the proposed deep learning algorithm in relation to the inertial sensor of the wearable device. For the six common activities in the UCI dataset and the data recorded, the recognition rates in the training sample reached 98.93% and 97.19%, respectively, and the recognition rates in the testing sample were 95.99% and 93.77%, respectively.

INDEX TERMS Wearable device, human activity recognition (HAR), inertial sensor, deep learning, classifier, convolutional neural network (CNN).

I. INTRODUCTION

Following the rapid development of computers and embedded systems, human activity recognition (HAR) through wearable devices and low-cost sensors has become an integral part of people's daily lives and is widely applied to various common domains including health management, medical monitoring, action recognition, rehabilitation activities, and remote control [1]–[10]. Wearable devices combining embedded sys-

tems and inertial sensors have been developed for activity recognition and are used in daily life and sports activities. The advantage of inertial sensors combined with embedded systems in wearable devices for motion monitoring and recognition is that no external environmental sensors such as radars, cameras, or infrared sensors are required for these wearable devices [11]–[13].

In addition, with their tiny size, lightness, low cost, and diminished power consumption, inertial sensors in wearable devices provide a solution for activity recognition in sports. According to Khalifa *et al.*, kinetic energy harvesting (KEH)

The associate editor coordinating the review of this manuscript and approving it for publication was Chong Leong Gan.

may help combat battery issues in wearable devices; KEH is mainly utilized as a generator and an HAR sensor, reducing power consumption by the sensor. The results indicated that Human Activity Recognition from Kinetic Energy could reduce overall system power consumption by 79% [5]. Some data were collected with professional equipment; in common daily activities, such equipment is inconvenient and expensive [14]. At present, smartphones have become an integral part of people's daily lives worldwide. Smartphones provide several functions in addition to basic telephone functions, such as diverse sensors integrated into the phones. Attempts have been made to use cell phone data for human action recognition [15], [16]. Thus, an increasing number of research has focused on the application of sensors in cell phones. Human activities were classified using datasets from smartphone devices [17].

Datasets that have been used for research and verification include the Human Activity Recognition Using Smartphones Dataset of UCI (University of California, Irvine) used in HAR research. Jain et al. proposed using a histogram of gradients and Fourier descriptors for feature extraction, and they applied them to the UCI HAR database. They proposed a support vector machine and k-nearest neighbor algorithm for classification [18]. Although the method they proposed had favorable performance, during feature extraction, they needed to conduct special feature engineering for each database. Gaikwad et al. proposed an HAR system specifically for smart military wearable devices. Through multilayer perceptron, the system classified activities. Their design obtained 270 ns classification time and consumed 120 mW of power. This system also utilized the UCI HAR database for verification and development [19]. Tufek et al. used an accelerometer and gyroscope to develop an activity recognition system. They used a three-layer long short-term memory (LSTM) model that reached 93.7% accuracy with the UCI HAR database. With additional data, the accuracy reached 97.4% [20]. In the activity recognition method of Sukor et al., who used acceleration sensors in smartphones, data were collected by adopting the acceleration of the open dataset using the data as the original input signal. Principal Component Analysis (PCA) was used to reduce the number of dimensions of the features and to extract information on the features of human activities in the time domain and the frequency domain for classification; six actions were recognized: standing, sitting, lying, walking upstairs, walking downstairs, and walking. The research method mainly involved partitioning the datasets into two parts, with 70% used for the training data and 30% for the test data; the experimental results indicated the accuracies to be 6.11% and 92.10%, respectively [21]. The aforementioned studies reported favorable performance. However, when conducting preprocessing, feature extraction requires specific feature engineering on each database, and after dimensionality reduction in deep learning, features cannot be completely retained, resulting in reduced accuracy. In the aforementioned literature, the UCI HAR dataset was employed for HAR

research. Thus, this study used the UCI HAR dataset to verify the proposed model and classify the data this study recorded using the same model.

Deep learning algorithms are currently widely researched. Unlike traditional machine learning methods requiring manual feature extraction, these algorithms can perform automatic feature extraction. Thus, in a series of studies, measured data were collected from sensing devices, the results were analyzed, and effective HAR systems were developed. Aljarrah et al. proposed a principle component analysis–bidirectional long short term memory approach to train bidirectional long short term memory recursive neural network models for predicting the identification of activities performed by participants in datasets; the number of dimensions of the dataset of 12 activities was reduced using PCA, achieving an accuracy rate of 97.64% [22]. Lee et al. proposed a method based on 1D convolutional neural networks (CNNs); 3-axis accelerometers in users' smartphones were collected to identify walking, running, and no motions. The acceleration data of x, y, and z axes served as the input for neural network, and the accuracy reached 92.7% [23]. For the feature learning method proposed by Zebin et al., the numerical values of the accelerometer and the gyro sensor served as the input, and feature learning was conducted automatically through CNNs; compared with SVM and multilayer perceptron methods, CNN demonstrated significant capabilities in terms of both computational complexity and classification accuracy [24]. Xu et al. created a CNN; the data from the 3D raw accelerometer could be directly used as the input for CNN training without any complex preprocessing required. The accuracy reached 91.97%, surpassing that of conventional SVM by 9% [25]. Zhang et al. proposed a new method using the attention mechanism of CNNs and HAR and concentrating attention into a multihead CNN, thereby facilitating feature extraction and selection and elevating accuracy to 95.4% [26]. Zebin et al. proposed a deep CNN model to classify five daily activities: walking, walking upstairs, walking downstairs, sitting for long periods of time and sleeping; the raw data from the accelerometer and the gyroscope of the wearable device served as the input, and the accuracy was 96.4% [27]. Xia et al. proposed a deep neural network combining the convolutional layer and LSTM; the model automatically performed feature extraction and classification without requiring manual feature extraction in conventional recognition approaches, and the accuracies for three open datasets (UCI, WISDM, and OPPORTUNITY) were 95.78%, 95.85%, and 92.63%, respectively [28]. Kańtoch indicated the correlation of sedentary lifestyle with increased pathogenic risk; seven healthy volunteers performed routine activities—sitting, walking, standing and squats—and the overall accuracy reached 82% [29]. Based on the aforementioned literature review, in employing deep learning algorithms, most neural networks automatically searched for and captured features in end-to-end networks without having to employ the traditional manual feature extraction method.

According to the aforementioned literature review, a CNN automatically extracting features in combination with wearable inertial sensors was developed to increase the accuracy of daily activity recognition; the proposed wearable device was placed on the waists of the experimental participants to measure motion signals and record the movements involved in daily activities. With the collected data, everyday activities could be recognized using an activity recognition algorithm. A CNN was adopted for the core of the main algorithm to conduct automatic feature extraction as well as classification and recognition of six everyday movements.

The remaining part of this study is arranged as follows: Section II presents the open dataset, the experimental participants' demographics, and characteristics of the wearable inertial sensor this study proposed, as well as the respective hardware structure. Section III introduces the activity recognition algorithm, including motion signal recording, signal normalization, data measurement, and format, and the model structure on the basis of CNN. Section IV shows the experimental results and discussion. Finally, Section V provides the conclusions.

II. EXPERIMENTAL SETUP
A. THE OPEN DATASET
In this study, a large and open dataset was required to verify the basis for the motion recognition learning model. Thus, the experimental data from the UCI dataset was adopted. The experimental participants carried Android smartphones (Samsung Galaxy S2) when performing certain everyday activities. The data in the experiment were derived from the Human Activity Recognition Using Smartphones Data Set. In these experiments, 30 healthy participants (aged between 19 and 48 years) volunteered to record data. Data were collected during six activities (walking, walking upstairs, walking downstairs, sitting, standing, and lying). These volunteers performed the six aforementioned activities with smartphones mounted on their waists. Sensor signals from the accelerometer and the gyroscope were analyzed, and three-axial acceleration and three-axial angular velocity were captured at a constant rate of 50 Hz. The dataset comprised data on the participants' movements obtained through camera recordings of the experiment with manual labeling. The UCI Machine Learning Repository supervised the data to guarantee quality. A total of 30 participants were featured in the dataset, the total number of entries was 10 299; these were randomly partitioned into two sets, where 70% of the participants, who were associated with 7352 entries, were selected and used in the training data; 20% of the training data (1470 entries) was used to test model accuracy; 30% of the participants, for whom the number of entries obtained was 2947, were used in the test data; multiple frames were used for each participant; the width of a frame was 256 signals, and sampling was performed in fixed-width sliding windows of 2.56 s and 50% overlap (128 readings/window); the signals of the accelerometer, and the numerical values for the three axes of the gyroscope signals were recorded every 0.02 s [30].

TABLE 1. Activity recognition experimental participants.

Parameters	Daily Activity
Men	21
Age(years)	22 ± 2
Height(cm)	165 ± 15
Weight(kg)	65 ± 15

B. DATA RECORDED IN THIS STUDY
1) PARTICIPANTS
1) This study collected data from motion signals produced by movements during everyday activities performed by 21 healthy participants (21 men; aged 22±2 years; height = 165±1 5 cm; weight = 65±15 kg); 2) the demographics and data for the experimental participants, for whom six activities were identified, are summarized in Table 1. In addition, in the motion-recognition experiment, 3) each of these 21 healthy participants was asked to mount an inertial sensing element on his or her waist and perform six everyday movements (walking, walking upstairs, walking downstairs, sitting, standing, and lying). Regarding the wearable inertia sensor (accelerometer and gyroscope), we converted the data obtained from the accelerator with 16,384 LSB/mg and converted the data obtained from the gyroscope with 131 LSB/degree. We then placed the inertia sensor at the center of the participant's waist. The wearing direction and location of the inertia sensor were strictly regulated. The *x*-axis is the righthand side of the body, the *y*-axis is the down side of the body, and the *z*-axis is the front of the body. This regulation was used throughout the experiment. Therefore, we used the fixed directions and location to conduct calibration. During the data-collection process, we had to ensure that the sensor-wearing location and directions were fixed. If the sensor was loosened during participants' activities, then the data recorded may be less accurate due to different directions. For this research, 4) most data were directional, and changes in direction resulted in decreased accuracy. Thus, cloth was place on the inertial sensing device to avoid imbalance in the inertial sensor; a belt was used to fasten the inertial sensing device and reduce vibration and effectively enhance the stability of the inertial sensor. 5) The data from Raspberry Pi of the microcontroller were transmitted to the cloud and downloaded to the computer with WiFi technology; the creation of the training and testing models was commenced. 6) The data on three-axis acceleration and three-axis angular velocity were collected at a constant rate of 50 Hz. The dataset obtained comprised 13 860 entries; each entry had 900 features, and each activity was performed for 15 minutes. Data were respectively sampled from 21 healthy participants; each one was asked to perform the movements involved in six everyday activities. 7) The collected motion data were partitioned into two sets respectively used in the training data and the test data, where 70% (9702 entries) were used in the training data, 20% of the training data (1940 entries) were used to test model accuracy, and 30% of the data from the total (4158 entries) were used to test model data.
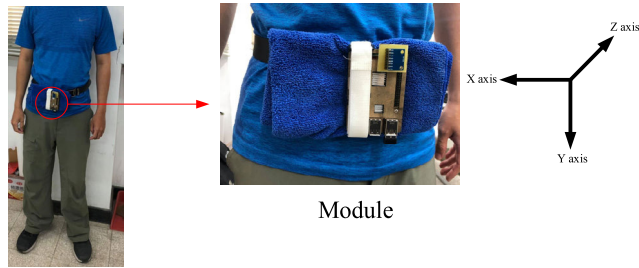
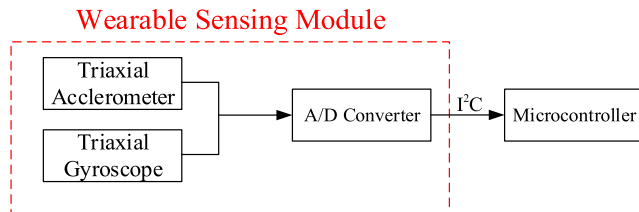**FIGURE 1.** Wearable sensor equipped on the waist of a participant.

FIGURE 2. Schematics of the wearable sensor.

FIGURE 3. Hardware structure of the wearable sensor.

FIGURE 4. The proposed activity recognition algorithm.

2) APPARATUS

The wearable inertial sensor employed in this study consisted of a 3-axis accelerometer and a 3-axis gyro. The sensor was equipped on the waist of each participant to record signals of his daily activities (Figure 1). The structure of the wearable device comprised an embedded microcontroller (Raspberry Pi 3), a 6-axis inertial sensor (MPU-6050), and a power supply (Figure 2). The microcontroller comprised a Broadcom BCM2387 processor, a 1.2-GHz 4-core ARM Cortex-A53.802.11 structure, and a 2-core Video Core IVR multimedia coprocessor as its graphics processing unit and featured a 32-G memory capacity and a size of 85 × 56 × 17 mm; its power supply connected through a micro USB slot with voltage and current of 5 V and 2.5 A, respectively. The microcontroller collected the numeric signals of human activities from the inertial sensor through the inter-integrated circuit (I^2C) and transmitted them to Raspberry Pi 3. The sensor consisted of a 3-axis accelerometer, a 3-axis gyro, and a 16-bit analog-to-digital converter. The sensor system was used to simultaneously collect the daily activities of human bodies, the resultant accelerations and angular velocities, and the three-dimensional activity spaces of the activities; the system output numeric signals of these activities. The accelerometer detected the acceleration of the inertial sensor on the waist of each participant in the *x* (right), *y* (downward), and *z* (frontal) directions in each test. The ranges of acceleration in these activities were ±2, ±4, ±8, and ±16 g; the ranges of angular velocities of said activities as detected through the gyro were ±250, ±500, ±1000, and ±2000 degrees per second. During the tests, the range and sensitivity of the accelerometer were set to ±16 g and 2048 LSB/g, respectively; those of the gyro were set to +2000°/s and 16.4 LSB/g, respectively. The output signal sampling frequencies of both the accelerometer and the gyro were set to 50 Hz. The activity recognition device had a mobile power supply that supplied direct current at 5 V
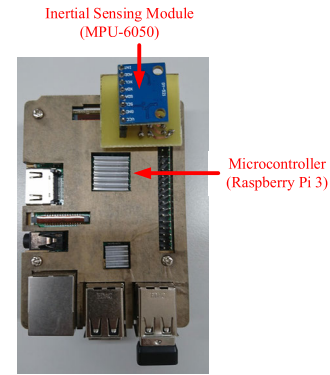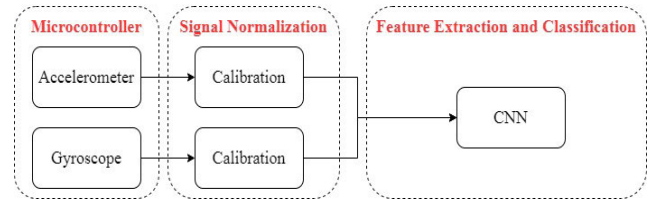
and 2.1 A. Figure 3 illustrates the hardware components of the wearable device, namely the microcontroller (Raspberry Pi 3) and 6-axis inertial sensor (MPU-6050).

III. ACTIVITY RECOGNITION ALGORITHM

The activity recognition algorithm procedures included collection of signals during activities and signal normalization. A deep learning algorithm and a CNN were employed to automatically enact feature extraction for recognizing human movements from everyday life. Fig. 4 shows the experimental flow for the proposed activity recognition algorithm. The process of the proposed activity recognition algorithm is as follows:

A. MOTION SIGNAL ACQUISITION

To obtain experimental data on acceleration and the angular velocity, the wearable inertial sensor was mounted onto the waists of 21 healthy experimental participants for testing six activities (walking, walking upstairs, walking downstairs, sitting, standing, and lying).

B. SIGNAL NORMALIZATION

Because the numerical precision of the accelerometer and of the gyroscope differed in the wearable inertial sensor, the uniform precision of the numbers had to be calibrated first to seven decimal places and represented with scientific notation to unify the lengths of the acceleration and the gyroscope of the data volume for each entry, ensure the preprocessed data, and avoid data inputting errors in performing the algorithm.

The key of our algorithm is the identification capability of the values of the six-axis accelerometer. The unit of the
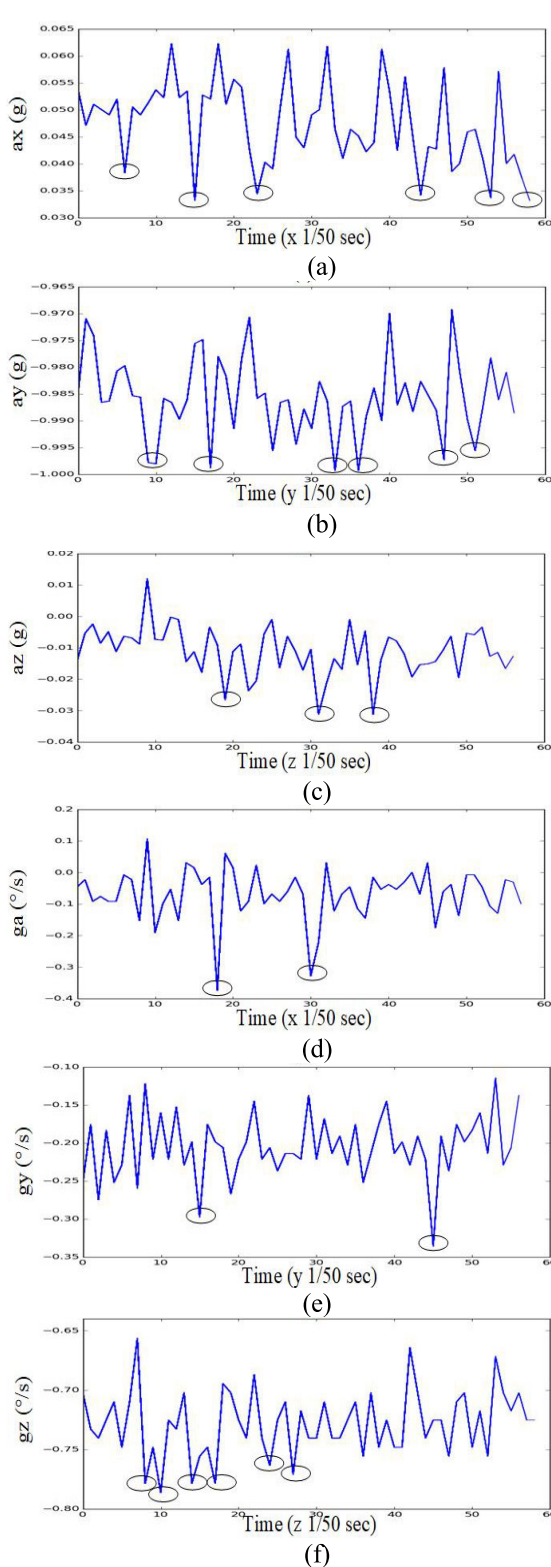
FIGURE 5. X, y, and z graphs for acceleration: (a); (b); (c) and gyroscope: (d); (e); (f).
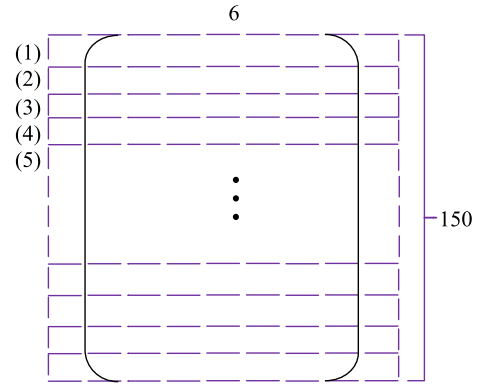
FIGURE 6. A dataset with scanning window which window size is 150.

that the az-axis exhibited substantial differences in the values. The value of walking az is approximately 0 g; that of going upstairs is below 0 g; and that of going downstairs is above 0 g. Sitting, standing, and lying down are three static postures. Sitting is similar to standing. The ay value of sitting is closer to 0 g, whereas that of standing is approximately -0.25 g. When an individual sits or lies down, their torso is horizontal. Therefore, their ax, ay, and az values change substantially. By using the differences in the figures, algorithms can be used to distinguish postures.

Observation of the x, y, and z axes for the three-axis acceleration and the gyroscope in Fig. 5 revealed that the waveforms of the three-axis acceleration and the gyroscope were generally aperiodic. The circled parts designate the most obvious features of the respective actions, namely the regional features of six movements: walking, walking upstairs, walking downstairs, sitting, standing, and lying to facilitate the subsequent classification with the classification algorithm.

C. DATA MEASUREMENT AND FORMAT

The sampling rate of the wearable inertial sensor was 50Hz. Thus, we have a matrix with the size of 150 to store data. Each matrix is arranged in $6 \times 150 = 900$ feature vectors; 6 stands for 6 dimensions, and 150 stands for the number of data. The first row of the vector for each element was then marked with the label for its type. The approach is illustrated in Fig. 6, and the same format and arrangement were required for both the training data and the test data to suit the input format of the classifier.

D. NETWORK STRUCTURE

CNN is an artificial neural network that uses the theory of deep learning, and comprises two parts: feature extraction and classification. The structure of the network system of the 1D CNN of this study consisted of four convolutional layers, three dropout layers, three fully connected layers, and a softmax layer, outputting the probability for each of the six activities. Our method differs from conventional cnn in that we do not include the pooling layers. Through experimentation, we discovered that although pooling layers reduce

y-axis is g, and that of the *x*-axis is the number of samples taken. Walking, ascending stairs, and descending stairs are dynamic activities. From these three activities, we observed
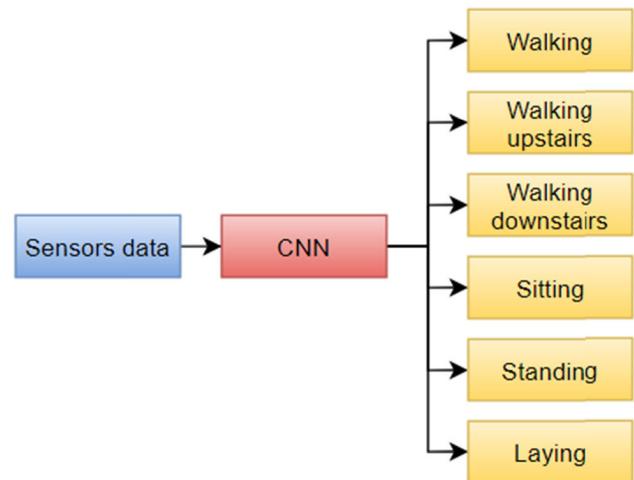
TABLE 2. CNN network structure.

Layer(type)
CONV1D-256
CONV1D-128
Dropout
CONV1D-64
CONV1D-32
Dropout
Flatten
Dropout
FC-256
FC-256
Softmax

FIGURE 7. Activity classification.

training time, they reduce accuracy as well. The reason for this may be that pooling layers reduced features based on the parameter setting. Therefore, we removed the pooling layers step to obtain more comprehensive features and the results show that the method is better used in har data.

1. Input: the six-axis data collected by the accelerometer and the gyroscope
2. Convolution: the convolution algorithm was employed to replace the matrix multiplication algorithm in traditional neural networks; the numbers of the kernels used in this study were 256, 128, 64, and 32; the sizes of their strides were all 1.
3. D ropout: to avoid overfitting in the neural network, the dropout in the experiment was set to 0.5
4. Activation Function: the rectified linear unit is the most widely used activation function in CNNS. The additional nonlinear relationships among the layers of the neural network converted some of the a neurons output into 0, leading to sparsity in the network, which could somewhat alleviate overfitting and the probability of gradients and divergence.
5. Output: The softmax layer was positioned as the output layer of the fully connected layer. The probability of each activity was computed by each unit (or node) in the softmax layer; the one with the highest probability was selected as the prediction.
6. Optimizer and learning rate: The adam optimizer was adopted, and the learning rate of the CNN model of this study was set to 0.00001.
7. Loss Function: For the categorical_crossentropy used in the model of this study, the closer the predictive value and the actual value were, the smaller the loss function became; by contrast, larger differences affected loss function more substantially.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The experiment was performed in a personal computer with Microsoft Windows 10, an Intel® Core Processor i7-8700, 32 GB RAM and GPU of NVIDIA GeForce RTX 2060. The performance proposed activity recognition algorithm was verified through cross-validation, which involved equations for accuracy, precision, recall, and error rate.

$$\text{Accuracy}(\%) = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \times 100\% \quad (1)$$

$$\text{Precision}_i(\%) = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \times 100\% \quad (2)$$

$$\text{P}_{macro}(\%) = \frac{1}{n} \sum_{i=1}^{n} \text{Precision}_i \times 100\% \quad (3)$$

$$\text{Recall}_i(\%) = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \times 100\% \quad (4)$$

$$\text{R}_{macro}(\%) = \frac{1}{n} \sum_{i=1}^{n} \text{Recall}_i \times 100\% \quad (5)$$

$$\text{F1-Score}(\%) = 2 \times \frac{\text{P}_{macro} \times \text{R}_{macro}}{\text{P}_{macro} + \text{R}_{macro}} \times 100\% \quad (6)$$

where i is one kind of 6 activities, TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

Precision refers to the degree of accuracy in the prediction of correct classification results (assuming that walking is positive and other activities are negative). Correct classification of walking then divided by correct classification plus wrong classification of walking, i.e., TP/TP+FP, shows how many of these predictions are correct. Recall means the correct opportunity of being predicted as correct under the correct classification (assuming that walking is positive and other activities are negative). Then, correct classification of walking divided by correct classification of walking plus wrong classification as other activities, i.e., TP/TP+FN, shows how many correct activities were identified. Simply put, a high precision means a high probability of finding a correct result,

TABLE 3. Evaluation indexes of the open dataset.

Accuracy	Precision	Recall	F1-Score
95.99%	96.04%	95.98%	96.01%

and a high recall means a high probability of finding more comprehensively. The purpose of F1-Score is to assess the balance between precision and recall. This index is only high when precision and recall are similar. If one of them performs well but the other poorly, then the F1 value will be small.

This paper adopted macro average precision (P*macro*), which first obtains the statistical index value of each class and then finds the mean of each class (assuming walking is positive, and all other activities are negative). After obtaining the precision and recall of a motion, it moves on to obtain those of the next motion. After six repetitions, we obtain six precisions and six recalls. Then the means of precisions and recalls were separately obtained, thereby obtaining the precision and recall of the multi-class P*macro* index.

A. EVALUATION INDEXES OF THE OPEN DATASET

In the experiment, each individual performed six movements involved in everyday activities in a laboratorial environment: walking, walking upstairs, walking downstairs, sitting, standing, and lying down. This study used a CNN as the method for feature extraction and classification. In using the open dataset, the accuracy reached 95.72%, the precision was 95.88%, the recall was 95.61%, and the F1-Score was 95.74%, as demonstrated in Table 3.

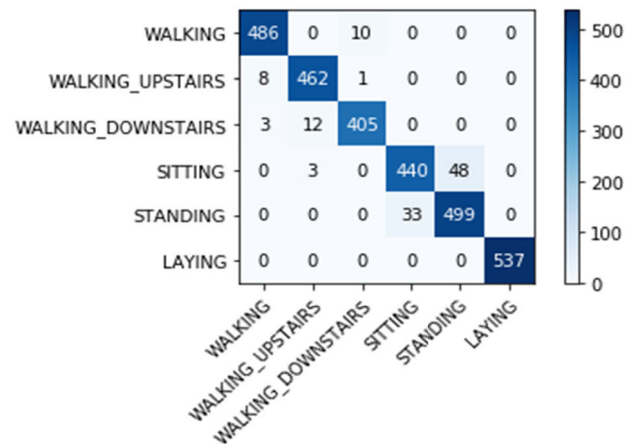## B. CONFUSION MATRIX OF THE OPEN DATASET

Fig. 8 shows that accurate classification is achieved in walking, walking upstairs, walking downstairs, and lying down; 10 data entries for sitting were classified as standing and 11 entries as sitting; sitting, standing, and lying down were classified as motionless actions. The confusion matrix indicates that lying down in the static actions was not wrongly classified. The reason for this is that significant changes in the numerical value of the gyroscope measurement occur during lying down; sitting and standing were susceptible to wrong classification in the model because the respective data were more similar, and no significant change was observed in the inertial sensor.

C. ACCURACY AND LOSS FUNCTION OF THE MODEL FOR THE OPEN DATASET

Fig. 9 and Fig. 10 demonstrate that, after 25-epoch training, the accuracy of the validation set and of the model both exceeded 97%, whereas the loss function of the validation set and of the model both decreased significantly.

D. EVALUATION INDEXES OF DATA THIS STUDY RECORDED

In the experiment, each subject performed six movements involved in everyday activities in a laboratorial environment:

FIGURE 8. Confusion matrix of the open dataset.
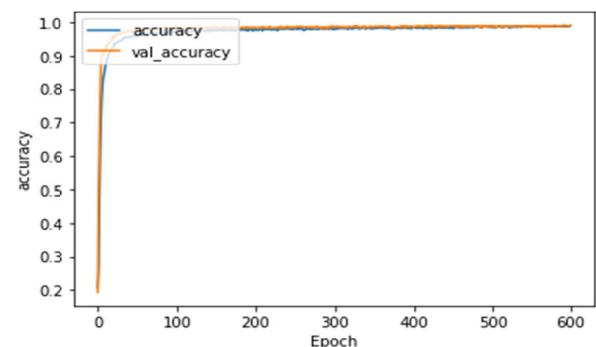
FIGURE 9. Model accuracy for open dataset.
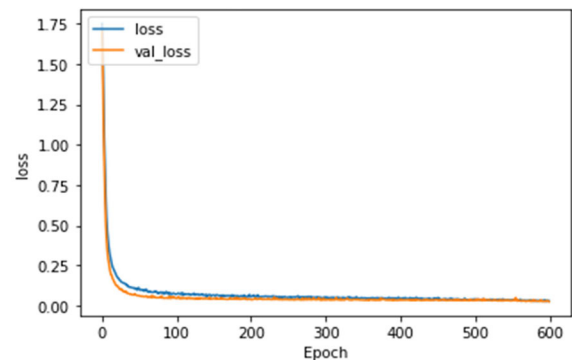
FIGURE 10. Loss function for model of open dataset.

TABLE 4. Evaluation indexes of data this study recorded.

Accuracy	Precision	Recall	F1-Score
93.77%	93.82%	93.82%	93.82%

walking, walking upstairs, walking downstairs, sitting, standing, and lying. This study used a CNN for feature extraction and classification. In using the recorded dataset, the accuracy reached 93.77%, the precision was 93.82%, the recall was 93.82%, and the F1-Score was 93.82, as shown in Table 4.
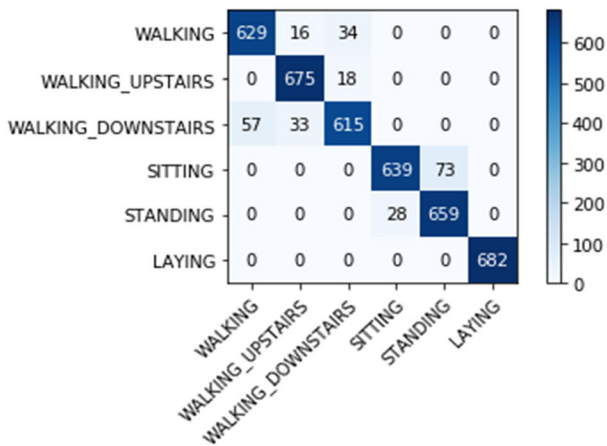
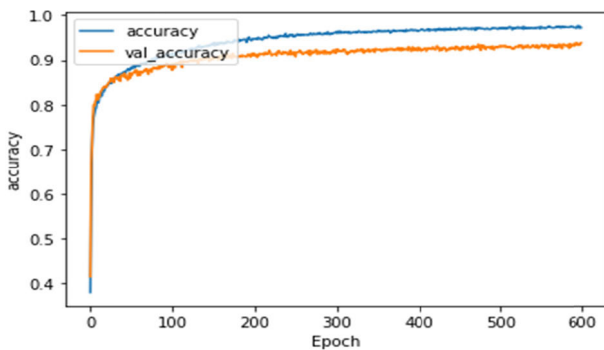**FIGURE 11.** Confusion matrix for data this study recorded.

FIGURE 12. Accuracy for model of data this study recorded.

FIGURE 13. Loss function for model of data this study recorded.

TABLE 5. The comparisons of several research models of the open dataset.

model	Accuracy
3-layer LSTM [20]	93.70%
LSTM-CNN [28]	95.78%
Bidir-LSTM [31]	93.79%
EHARS [32]	93.92%
CNN-LSTM [33]	92.13%
CNN-LSTM [34]	93.40%
Ours	95.99%

E. CONFUSION MATRIX OF DATA THIS STUDY RECORDED

Fig. 11 shows that all classifications for lying down were correct. The reason for this is that significant change in the numerical value on the gyroscope occurs when someone lies down. With the remaining data, the accuracy was not observed to be higher than that for the open dataset in terms of classifying activities involving movement. This was because the data this study recorded had not undergone filtering and manual processing, relying completely on the CNN for feature extraction, leading to a lower accuracy in comparison with that of the open dataset, but nonetheless achieving accuracy of 93.77%.

F. ACCURACY AND LOSS FUNCTION OF THE MODEL FOR DATA THIS STUDY RECORDED

Fig. 12 and Fig. 13 demonstrate that, after 350-epoch training, the accuracy of the validation set and of the model both become more stable. The public databases underwent filtering and manual handling, whereas the data we collected were only normalized. Under the same CNN network framework, we observed that data recorded in different environments exhibited the same trend. Therefore, the modified CNN algorithm we proposed was effectively verified.
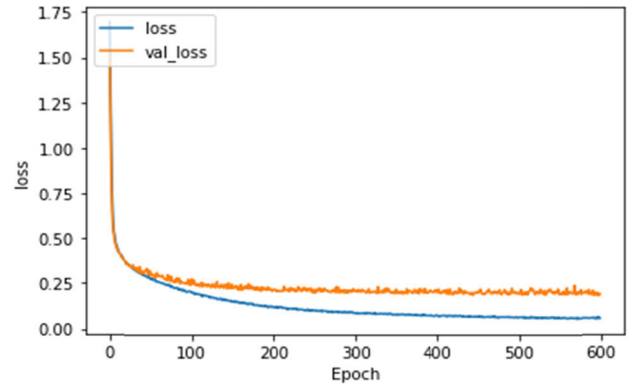
G. K-FOLD CROSS-VALIDATION IN BOTH OPEN DATASET AND DATA THIS STUDY RECORDED

We utilized k-fold cross-validation by dividing all data into 10 equal parts and letting $k = 10$ to assess the algorithm. We used $k - 1$ for training and the deducted 1 portion for testing. We ensured that the testing data were not used in training. After 10 repetitions, the average accuracy of the open database was 95.08%, and the mean accuracy of the database we recorded was 87.88%.

H. THE COMPARISONS OF SEVERAL MODELS OF THE OPEN DATASET

The Table 5 lists the accuracies of several researches by using UCI open data set, we can find the proposed algorithm is better than other researches. Reference [20] used a three-layer LSTM model. Although reference [28] has an accuracy similar to ours, like references [33] and [34] it did not use only CNN but also LSTM, and thus the complexity of the algorithm was increased. This study used only CNN and could obtain similar or even superior accuracy. Reference [31] used bidirectional LSTM, and it took 50,000 iterations to converge. By contrast, the present study only required 600 iterations to obtain superior accuracy. Reference [32], like us, used CNN, but we differed from it in that after research we excluded the pooling layers and obtained higher accuracy.

V. CONCLUSION

This study proposed an activity classification algorithm using deep learning and a 1D CNN. The wearable inertial

sensing device comprised the hardware, and the software used an activity recognition algorithm. The activity recognition algorithm harvested data from motion signals, detected by the sensor, for three-axis acceleration and three-axis angular velocity. This was followed by signal normalization. Features were then automatically extracted through the convolutional layer of a CNN and classification performed through the softmax layers to identify six everyday activities-walking, walking upstairs, walking downstairs, sitting, standing and lying-performed by 21 participants in the open dataset and the recorded data. The overall accuracies of the open dataset in the training data and the testing data were 98.93% and 95.99%, respectively; the overall accuracies of the data this study recorded in the training data and the testing data were 97.19% and 93.77%, respectively. The experimental results successfully verified that the proposed CNN could be considered an effective method based on inertial sensors in recognizing everyday human activities and can be used in the future to evaluate the amount of rehabilitation exercise of individuals with reduced mobility, such as patients undergoing dialysis, thereby affirming that the algorithm we proposed is feasible.

REFERENCES

[1] 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.

[2] M. Seiffert, F. Holstein, R. Schlosser, and J. Schiller, "Next generation cooperative wearables: Generalized activity assessment computed fully distributed within a wireless body area network," *IEEE Access*, vol. 5, pp. 16793–16807, 2017.

[3] Y. Chen and C. Shen, "Performance analysis of smartphone-sensor behavior for human activity recognition," *IEEE Access*, vol. 5, pp. 3095–3110, 2017.

[4] M. Janidarmian, A. Roshan Fekr, K. Radecka, and Z. Zilic, "A comprehensive analysis on wearable acceleration sensors in human activity recognition," *Sensors*, vol. 17, no. 3, p. 529, Mar. 2017.

[5] S. Khalifa, G. Lan, M. Hassan, A. Seneviratne, and S. K. Das, "HARKE: Human activity recognition from kinetic energy harvesting data in wearable devices," *IEEE Trans. Mobile Comput.*, vol. 17, no. 6, pp. 1353–1368, Jun. 2018.

[6] 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.

[7] 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.

[8] M. Ueda, H. Negoro, Y. Kurihara, and K. Watanabe, "Measurement of angular motion in golf swing by a local sensor at the grip end of a golf club," *IEEE Trans. Human-Machine Syst.*, vol. 43, no. 4, pp. 398–404, Jul. 2013.

[9] L. Cantelli, G. Muscato, M. Nunnari, and D. Spina, "A joint-angle estimation method for industrial manipulators using inertial sensors," *IEEE/ASME Trans. Mechatronics*, vol. 20, no. 5, pp. 2486–2495, Oct. 2015.

[10] N. Tufek and O. Ozkaya, "A comparative research on human activity recognition using deep learning," in *Proc. 27th Signal Process. Commun. Appl. Conf. (SIU)*, Sivas, Turkey, Apr. 2019, pp. 1–4.

[11] 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.

[12] I. C. Gyllensten and A. G. Bonomi, "Identifying types of physical activity with a single accelerometer: Evaluating laboratory-trained algorithms in daily life," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2656–2663, Sep. 2011.

[13] A. Wang, G. Chen, J. Yang, S. Zhao, and C.-Y. Chang, "A comparative study on human activity recognition using inertial sensors in a smartphone," *IEEE Sensors J.*, vol. 16, no. 11, pp. 4566–4578, Jun. 2016.

[14] R. Poppe, "Vision-based human motion analysis: An overview," *Comput. Vis. Image Understand.*, vol. 108, nos. 1–2, pp. 4–18, 2007.

[15] J.-L. Reyes-Ortiz, L. Oneto, A. Ghio, A. Sama, D. Anguita, and X. Parra, "Human activity recognition on smartphones with awareness of basic activities and postural transitions," in *Artificial Neural Networks and Machine Learning—ICANN*. Cham, Switzerland: Springer, 2014, pp. 177–184.

[16] N. Davies, D. P. Siewiorek, and R. Sukthankar, "Activity-based computing," *IEEE Pervas. Comput.*, vol. 7, no. 2, pp. 58–61, Apr. 2008.

[17] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proc. AAAI*, vol. 5, 2005, pp. 1541–1546.

[18] A. Jain and V. Kanhangad, "Human activity classification in smartphones using accelerometer and gyroscope sensors," *IEEE Sensors J.*, vol. 18, no. 3, pp. 1169–1177, Feb. 2018.

[19] N. B. Gaikwad, V. Tiwari, A. Keskar, and N. C. Shivaprakash, "Efficient FPGA implementation of multilayer perceptron for real-time human activity classification," *IEEE Access*, vol. 7, pp. 26696–26706, 2019.

[20] N. Tufek, M. Yalcin, M. Altintas, F. Kalaoglu, Y. Li, and S. K. Bahadir, "Human action recognition using deep learning methods on limited sensory data," *IEEE Sensors J.*, vol. 20, no. 6, pp. 3101–3112, Mar. 2020.

[21] A. S. A. Sukor, A. Zakaria, and N. A. Rahim, "Activity recognition using accelerometer sensor and machine learning classifiers," in *Proc. IEEE 14th Int. Colloq. Signal Process. Appl. (CSPA)*, Batu Feringghi, Mar. 2018, pp. 233–238.

[22] A. A. Aljarrah and A. H. Ali, "Human activity recognition using PCA and BiLSTM recurrent neural networks," in *Proc. 2nd Int. Conf. Eng. Technol. its Appl. (IICETA)*, Al-Najef, Iraq, Aug. 2019, pp. 156–160.

[23] S.-M. Lee, S. Min Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Jeju, Island, Feb. 2017, pp. 131–134.

[24] T. Zebin, P. J. Scully, and K. B. Ozanyan, "Human activity recognition with inertial sensors using a deep learning approach," in *Proc. IEEE Sensors*, Orlando, FL, USA, Oct. 2016, pp. 1–3.

[25] W. Xu, Y. Pang, Y. Yang, and Y. Liu, "Human activity recognition based on convolutional neural network," in *Proc. 24th Int. Conf. Pattern Recognit. (ICPR)*, Beijing, China, Aug. 2018, pp. 165–170.

[26] H. Zhang, Z. Xiao, J. Wang, F. Li, and E. Szczerbicki, "A novel IoT-perceptive human activity recognition (HAR) approach using multihead convolutional attention," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1072–1080, Feb. 2020.

[27] T. Zebin, P. J. Scully, N. Peek, A. J. Casson, and K. B. Ozanyan, "Design and implementation of a convolutional neural network on an edge computing smartphone for human activity recognition," *IEEE Access*, vol. 7, pp. 133509–133520, 2019.

[28] K. Xia, J. Huang, and H. Wang, "LSTM-CNN architecture for human activity recognition," *IEEE Access*, vol. 8, pp. 56855–56866, 2020.

[29] E. Kantoch, "Human activity recognition for physical rehabilitation using wearable sensors fusion and artificial neural networks," in *Proc. Comput. Cardiol. Conf. (CinC)*, Rennes, France, Sep. 2017, pp. 1–4.

[30] J. Reyes-Ortiz, D. Anguita, A. Ghio, L. Oneto and X. Parra. (Apr. 2012). *UCI Machine Learning Repository: Human Activity Recognition Using Smartphones Data Set*. [Online]. Available: https://archive.ics uci.edu/ml/datasets/human+activity+recognition+using+smartphones

[31] S. Yu and L. Qin, "Human activity recognition with smartphone inertial sensors using bidir-LSTM networks," in *Proc. 3rd Int. Conf. Mech., Control Comput. Eng. (ICMCCE)*, Huhhot, China, Sep. 2018, pp. 219–224.

[32] S. Dhanraj, S. De, and D. Dash, "Efficient smartphone-based human activity recognition using convolutional neural network," in *Proc. Int. Conf. Inf. Technol. (ICIT)*, Bhubaneswar, India, Dec. 2019, pp. 307–312.

[33] R. Mutegeki and D. S. Han, "A CNN-LSTM approach to human activity recognition," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Fukuoka, Japan, Feb. 2020, pp. 362–366.

[34] S. Deep and X. Zheng, "Hybrid model featuring CNN and LSTM architecture for human activity recognition on smartphone sensor data," in *Proc. 20th Int. Conf. Parallel Distrib. Comput., Appl. Technol. (PDCAT)*, Gold Coast, QLD, Australia, Dec. 2019, pp. 259–264.

CHIH-TA YEN (Member, IEEE) received the B.S. degree from the Department of Electrical Engineering, Tamkang University, Taiwan, in 1996, the M.S. degree from the Department of Electrical Engineering, National Taiwan Ocean University, Taiwan, in 2002, and the Ph.D. degree from the Department of Electrical Engineering, National Cheng Kung University, Taiwan, in 2008. He is currently a Professor of artificial intelligence applications, multiple access communications, and optical design technologies with the Department of Electrical Engineering, National Formosa University, Yunlin, Taiwan. His major research interests include multiuser optical communications, wireless communication systems, machine learning, deep learning, and optical design.

YI-KAI HUANG was born in Yunlin, Taiwan, in October 1996. He is currently pursuing the degree with the Department of Electrical Engineering, National Formosa University. His major research interest includes deep learning.

• • •

JIA-XIAN LIAO was born in Taichung, Taiwan, in September 1996. He is currently pursuing the degree in electrical engineering with National Formosa University, Taiwan. His major research interests include deep learning, machine learning, and embedded systems.