*Original Research Article*

DIGITAL HEALTH

DIGITAL HEALTH
Volume 10: 1–15

DOI: 10.1177/20552076231223804
journals.sagepub.com/home/dhj

Sage

# Clinical human activity recognition based on a wearable patch of combined tri-axial ACC and ECG sensors

**Yanling Ren[1], Minqi Liu[2], Ying Yang[1], Ling Mao[2] and Kai Chen[1]**

## Abstract

**Background:** In digital medicine, human activity recognition (HAR) can be used to track and assess a patient's progress throughout rehabilitation, enhancing the quality of life for the elderly and the disabled.

**Methods:** A patch-type flexible sensor that integrated dynamic electrocardiogram (ECG) and acceleration signal (ACC) was used to record the signals of the various behavioral activities of 20 healthy volunteers and 25 patients with pneumoconiosis. Seven HAR tasks were then carried out on the data using four different deep learning methods (CNN, LSTM, CNN-LSTM and GRU).

**Results:** When ECG and ACC were obtained simultaneously, the overall accuracy rates of HAR for healthy group were 0.9371, 0.8829, 0.9843 and 0.9486 by the CNN, LSTM, CNN-LSTM and GRU models, respectively. In contrast, the overall accuracy rates of HAR for the pneumoconiosis patients' group were 0.8850, 0.7975, 0.9425 and 0.8525 by the four corresponding models. The accuracy of HAR for both groups using all four models is higher than when only ACC signal is detected.

**Conclusion:** The addition of the ECG signal significantly improves HAR outcomes in the group of healthy individuals, while having relatively less enhancing effects on the group of patients with pneumoconiosis. When ECG and ACC signals were combined, the increase in HAR accuracy was notable compared to cases where no ECG data was provided. These results suggest that the combination of ACC and ECG data can represent a novel method for the clinical application of HAR.

## Keywords

Human activity recognition, tri-axial accelerometer (ACC), electrocardiogram (ECG), CNN-LSTM network

Submission date: 19 September 2023; Acceptance date: 14 December 2023

## Introduction

Human activity recognition (HAR) has important implications in health detection and self-monitoring.[1] The utilization of various sensors to monitor patient movements, track and analyze the patient's progress throughout the rehabilitation process is becoming increasingly vital. At present, computer vision and sensor technology are the main methods to recognize human behavior.[2] Compared with computer vision recognition, sensor recognition is more convenient, can better protect user privacy and is less susceptible to external influences in the course of usage. Wearable sensors have been applied in HAR for many years. For example, a single inertial measurement

[1]School of Mechanical Engineering, Hangzhou Dianzi University, Hangzhou, China
[2]Department of Pneumoconiosis, Shanghai Pulmonary Hospital, Tongji University, Shanghai, China

Yanling Ren and Minqi Liu have contributed equally to this work and share first authorship.

**Corresponding authors:**
Ling Mao, Department of Pneumoconiosis, Shanghai Pulmonary Hospital, Tongji University, Shanghai, China.
Email: drlingmao@163.com

Kai Chen, School of Mechanical Engineering, Hangzhou Dianzi University, Hangzhou, China.
Email: kchen@hdu.edu.cn

unit inserted at the ankle was utilized to monitor foot movement, evaluating the daily monitoring of the human body throughout therapy.[3] Fabien Massé et al.[4] reported a work that used inertial sensors and blood pressure monitors to record patients' everyday activities, including changes in posture. Lin et al.[5] used an accelerometer placed around the waist to capture behavioral activity from spinal cord injury patients, giving a tool to improve patient care by tracking activity changes in clinical clinics and at home. In addition to the aforementioned inertial or acceleration signals for activity identification, wearable sensors can also be employed to gather physiological signals, including the electrocardiogram (ECG) and breathing rate. For example, a wearable sensor for detection of the QRS waveform was designed to track ECG signals in the reported work.[6] Additionally, research has demonstrated that multifunctional body sensors may be used to continually monitor the ECG signal produced by the human body.[7] Chu et al.[8] used a wearable sensor put on the chest to gather chest and abdominal tension during breathing in order to assess respiratory rate. The physiological signals are significant data that reflect the vital qualities and activities of the human body. However, rarely works were reported to integrate ECG into HAR and put it into practical application.

Various machine learning methods such as SVM, decision trees and KNN have been extensively employed in previous studies for activity recognition utilizing accelerometer data collected by wearable sensors.[9–11] However, traditional machine learning algorithms mainly rely on manual labor to analyze and extract features from raw data.[12] As artificial intelligence technology advances, more researchers have been focusing on the application of deep learning algorithms for HAR.[13] For example, Lemieux et al.[14] and Murad et al.[15] built a 1D-CNN model and a RNN model for HAR, respectively, and reported that the CNN and RNN models had a better performance in classification than that of SVM and KNN. Moreover, models can also be combined with other one or more algorithms to optimize the process of HAR. For example, Zheng et al.[16] and Wang et al.[17] added additional mechanisms to a standard CNN model, and the traits extrapolated from the CNN were transferred to the new mechanism for further analysis in order to increase the model's accuracy. Attributed to the CNN's capacity for multi-scale feature extraction, the fused features can be conveyed into the GRU. Lu et al.[18] constructed a CNN–GRU model using global average pooling, achieving accuracy levels of 96.41, 96.67 and 96.25%, respectively. Using a deep model built on a recurrent neural network and attention-based two-way gated loop units, Mekruksavanich et al.[19] were able to achieve 99, 95.82 and 88.24% accuracy on the dataset. The performance of 1D-CNN was evaluated on the public datasets UCI, WISDM and PAMAP2. The findings were 98, 97.8 and 90.27%, respectively.[20] Li et al.[21] create a HAR algorithm (HAR-WCNN) based on wide-time CNN and multi-environment sensor data and achieved good results in CASAS. A spatial distance matrix based on the arrangement of environmental sensors was provided to reduce noise. Soni et al.[22] created a deep neural network (DNN) combining bidirectional long short-term memory (Bi-LSTM) and CNN, with an accuracy of 97.96 and 97.15% on WISDM and UCI-HAR, respectively.

The integration of physiological signals into behavioral analysis, such as the ECG, has been steadily gaining researchers' attention.[23–27] For instance, Pawar et al.[28] studied the classification of motion, arm movement, walking and climbing stairs through ECG signals using a single-lead wearable ECG device. ECG is also one of the necessary signals for emotion recognition, and it has a wide application in multi-task learning, such as game difficulty classification and tired driving detection.[29] Therefore, it is of clinic merit to simultaneously collect acceleration and ECG signals through wearable sensors to achieve the purpose of activity recognition. Pneumoconiosis, caused by inhaling dust into the lungs, is a systemic illness characterized by diffuse fibrosis of lung tissue.[30,31] Pneumoconiosis can result in dyspnea, which may lead to respiratory failure in severe cases. Consequently, HAR for patients with pneumoconiosis can be instrumental in monitoring their activities at home and assessing the effectiveness of medication.[32]

In this study, we examined the behavioral activities of the human body by integrating acceleration signals and ECG signals collected during various activities from 20 healthy individuals and 25 patients with pneumoconiosis, using a wearable patch. We propose a deep learning model of CNN–LSTM, for automatic feature extraction in HAR. Spatial features and spatiotemporal features were extracted by CNN and LSTM, respectively, and we analyzed the impact of the ECG signal on HAR.

In this study, we introduced physiological signaling ECG to HAR in this study and assessed it using several deep models. Not only can the human body's heart rate be monitored in real time, but the human body's behavior may also be supervised. The main contributions are as follows:

1. For data collection, a flexible wearable device with an integrated accelerometer and ECG was proposed, and seven basic behaviors of 20 healthy volunteers and 4 basic behaviors of 25 pneumoconiosis patients were gathered. Acceleration data and ECG data were obtained from 45 volunteers.
2. To categorize the collected datasets, a CNN–LSTM depth model is proposed, which can enable the automatic extraction of spatial and temporal information. The results from proposed CNN–LSTM model was compared to the results of other three deep models of CNN, LSTM and GRU, and it demonstrated that the

CNN-LSTM model had the best classification performance and stability.

3. To investigate the influence of ECG signals on HAR and compare the classification outcomes of the four models with only ACC signals and ACC + ECG signal inputs, with the conclusion that ECG can improve HAR. Clinical monitoring of patients' physical activity levels, heart health and overall physiological status is anticipated in the future of ECG.

## Materials and methods

### Participants

The data required for this experiment were obtained from a total of 45 volunteers, comprising 25 patients diagnosed with pneumoconiosis and 20 healthy young individuals. The pneumoconiosis patients were recruited from the Department of Pneumoconiosis at Shanghai Pulmonary Hospital between July and August 2023, while the healthy volunteers were recruited from Hangzhou Dianzi University during the same period. All participants signed informed consent forms after receiving a detailed description of the experiment. This study was conducted in accordance with the guidelines of the Helsinki Declaration of the World Medical Association (2000) and was approved and supervised by the Ethics Committee of Shanghai Pulmonary Hospital (approval No. L23-226). The demographic information of the participants is presented in Table 1.

The inclusion criteria for all participants were as follows: (a) Healthy young volunteers between the ages of 20 and 30 years, with no underlying medical conditions. (b) Patients with pneumoconiosis by a professional pneumoconiosis expert, aged between 40 and 80 years old. (c) Participants capable of understanding the given instructions.

The exclusion criteria were as follows: (a) Patients with liver and kidney insufficiency, as well as cardiac insufficiency. (b) Patients with other diseases that affect normal walking. (c) Individuals who experience allergic reactions to the wearable patch. (d) Patients with significant cognitive impairment who are unable to respond to the given instructions.

**Table 1.** The demographic information of the participants.

| Nature | Healthy group | Patients with pneumoconiosis |
|---|---|---|
| Number | 20 | 25 |
| Sex | Male: 15 | Male: 20 |
| | Female: 5 | Female: 5 |
| Age | 24(±1) | 65(±11) |
| Disease grade | N/A | I: 20<br>III: 5 |

### Device

The data for this experiment were collected using a flexible ECG Patch (by Vivalink, CA, USA), which features an embedded three-axis accelerometer and an embedded single-lead ECG sensor. The tri-axial accelerometer and the single-lead ECG sensor were used to record acceleration (ACC) signals and ECG signals, respectively. Using only one device to gather data on the left chest avoids the challenges of multi-sensor fusion, various data time calibration and the annoyance of sticking to the body caused by using numerous separate devices. The public data UCI, for example, use many sensors positioned on the subject's chest, right wrist and left ankle, as well as a lead ECG sensor, which will create significant disruption to the experimental collection.[33] As depicted in Figure 1, the ECG patch is characterized by its flexibility, portability and strong resistance to interference, making it easy to attach to the left chest skin of the volunteers. It can be worn continuously for 72 h, and the collected signals can be effortlessly downloaded to a personal computer. During the experiment, each volunteer affixed an ECG patch to their left chest, and both ACC and ECG signals were collected at a frequency of 128 Hz.

### Experimental procedures

We collected the data from 20 healthy young volunteers with seven basic behaviors, namely walking, jogging, sitting, standing, lying, going upstairs and going downstairs. The experiments were conducted at a corridor and a stairway in an office building. A special personnel

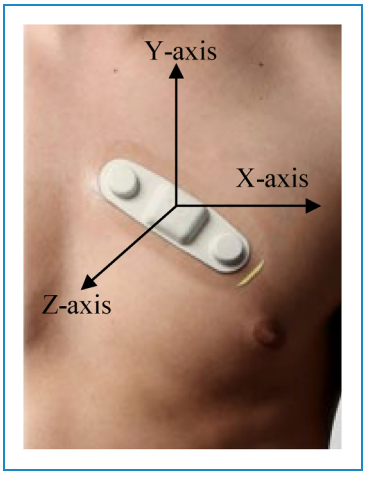


**Figure 1.** The graph of the ECG patch placed on the left chest of a subject.

recorded the experimental process during the experiment. Each behavior was recorded for a duration of 30 s, with participants adopting a natural posture and speed. There was a 1-min break between each behavioral data collection in order to lower the heart rate of volunteers. To visualize the characteristics of the ACC raw data for each axis, Figure 2 shows the acceleration waveform of a volunteer for 9 s per behavioral activity.

To ensure the safety of patients, the experiments on pneumoconiosis patients were carried out in the inpatient wards and outdoor walkway of Shanghai Pulmonary Hospital under the doctor's supervision. In addition, only four behaviors (sitting, standing, lying and walking) were collected due to the medical conditions of the patients. Volunteers performed each behavior for 30 s, adopting a natural posture and speed. After each behavior, a 3-min break was taken before proceeding to the next activity. In this experiment, all patients used the same type of wearable ECG patch. Figure 3 shows the ECG waveform of each behavioral activity for 6 s in a patient with pneumoconiosis.

Both the healthy volunteers and pneumoconiosis patients repeated the same experiment three times, with all data collection conducted in a quiet environment.

The samples and proportions of each behavior in the healthy human data and pneumoconiosis patients' data collected in this experiment are shown in Table 2.

## Data pre-processing

To enhance the accuracy of the Deep learning models, the raw ECG data and the raw ACC data undergoes the following preprocessing steps.

The ECG signal is a faint physiological signal characterized by non-linearity and non-stationarity.[34] The ECG signals captured by the ECG Patch can be corrupted by noise originating from various factors, including the acquisition instrument and the surrounding environment. Two primary sources of noise are baseline drift and myoelectric interference.[35] To solve these problems, we employed the discrete wavelet transform (DWT) to eliminate baseline drift in ECG signals and the Butterworth low-pass filter to eliminate EMG interference. DWT decomposes ECG signals into wavelet functions of varying scales, allowing the homeopathic frequency and amplitude features of the ECG to be examined on several time scales. For ACC signals, we use a conventional Butterworth filter to remove the introduced noise and accurately eliminate singular values.

Due to the variability in heart rate (HRV), there will be small fluctuations in the R-R interval of each person's heartbeat.[36] Normalization of the acquired data is needed to avoid the singular data. To improve the number of samples received by the CNN–LSTM model, the data was split into two 2-s sliding windows with 50% overlap.[37]

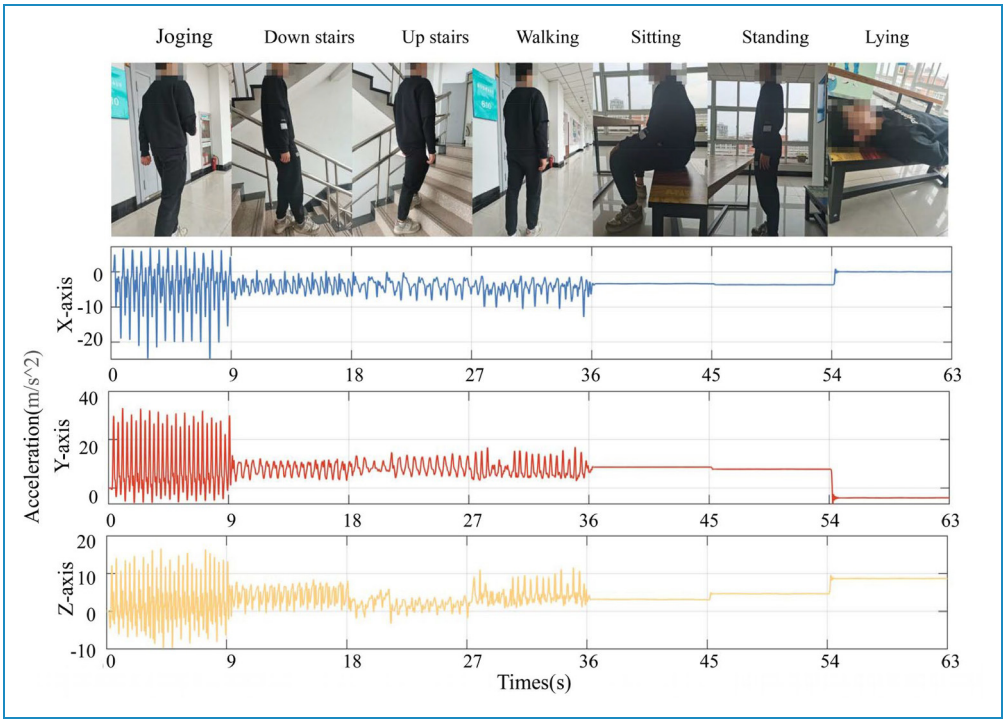


**Figure 2.** ACC signal waveforms of a young volunteer for seven activities, with each interval corresponding to one activity.

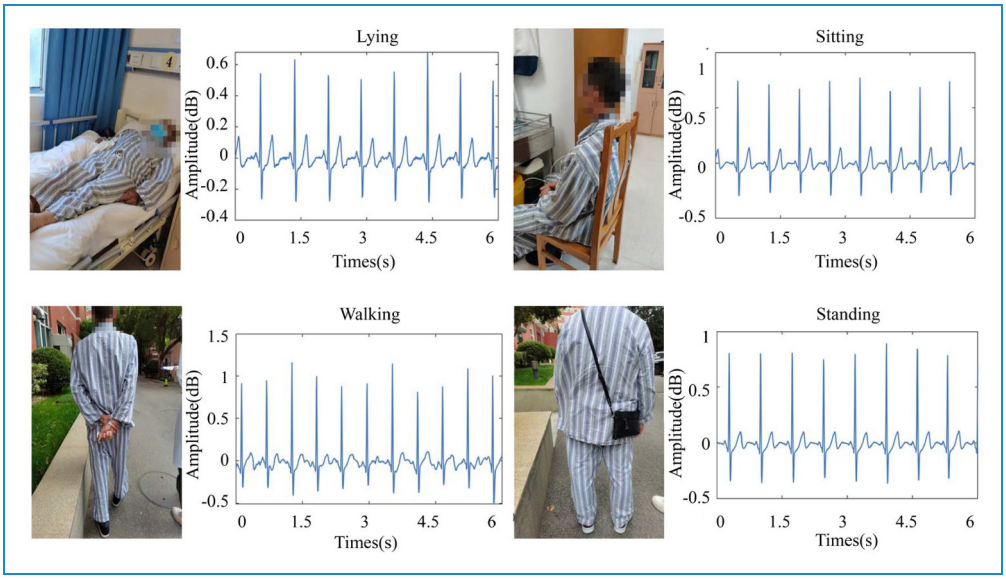


**Figure 3.** ECG signal waveforms of a patient with pneumoconiosis for four behavioral activities.

**Table 2.** Number and proportions of samples for each activity.

| Group | Activity | Number | Proportions (%) |
|---|---|---|---|
| Healthy Young Group | Sitting | 1820 | 15.97 |
| | Standing | 1820 | 15.97 |
| | Lying | 1820 | 15.97 |
| | Walking | 1820 | 15.97 |
| | Jogging | 1810 | 15.88 |
| | Going downstairs | 1242 | 10.90 |
| | Going upstairs | 1066 | 9.34 |
| Pneumoconiosis Patients Group | Sitting | 2150 | 25 |
| | Standing | 2150 | 25 |
| | Lying | 2150 | 25 |
| | Walking | 2150 | 25 |

## *Neural network model*

Figure 4 illustrates the structure of the CNN–LSTM model, consisting of seven layers. These layers collaborate to capture sequential data's temporal and geographical details. The preprocessed data was put into the model to extract the temporal data in order to derive spatial features. To reduce the possibility of overfitting, a discarded layer was added to the CNN–LSTM model.

In the CNN component, two convolutional layers were utilized, each equipped with a convolution kernel of different lengths to perform convolution operations and generate feature maps. The first layer comprises 64 filters, each with a size of 3. The second layer consists of 128 filters with a size of 5. Both convolutional layers employ the rectified linear unit (Relu) as their activation function.

The dimension of the features recovered by convolution was lowered, and redundant information was removed by the pooling layer, which was sandwiched between two convolutional layers.

LSTM is a distinct type of RNN that addresses for RNN's short-term memory limitations by incorporating gates, which provides an advantage when dealing with time series data.[38] The LSTM model is also capable of mitigating the issues of gradient vanishing and exploding during network training.[39] Consequently, in comparison to the conventional 1D-CNN model, the memory unit of the CNN–LSTM model can more effectively capture dependencies within sequential data, thereby enhancing the model's accuracy.

The output layer of the model was made up of two completely connected layers, each in charge of performing the final classification task. The activation function, which employed Softmax and generated a probability vector showing the likelihood of the current sample belonging to each activity, included seven neurons in the final fully connected layer.

TensorFlow 2.2 and Python 3.7 were applied to implement the proposed CNN–LSTM model. Cross-entropy was used as a loss function in model assessment to calculate the difference between the anticipated value and the actual value. Adam was used as an optimizer for parameter updates and calculations in model training, with a learning

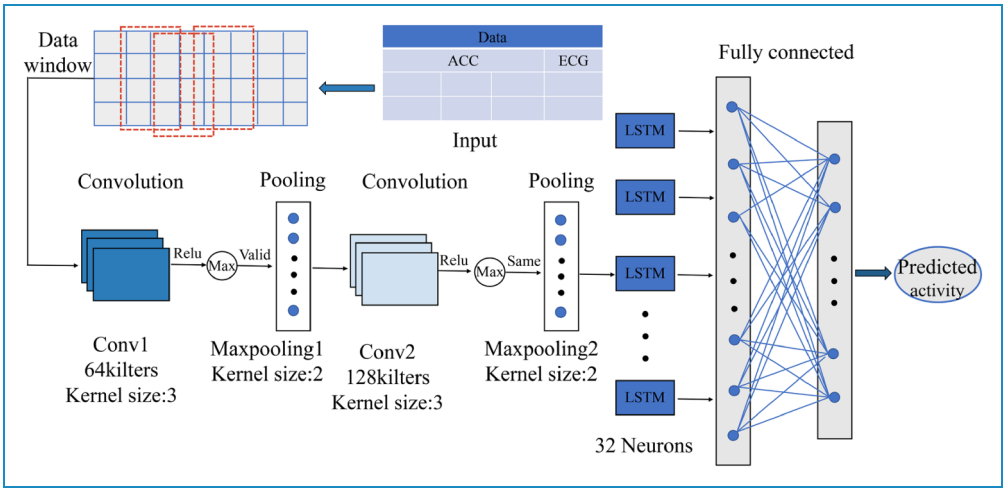


**Figure 4.** CNN-LSTM model structure.

**Table 3.** List of hyper-parameters of the CNN-LSTM model.

| Stage | | | | Selected values |
|---|---|---|---|---|
| Data preprocessing | Layer's name | Activation function | | |
| CNN | Convolution_1 | Relu | Filters | 64 |
| | | | Kernel size | 3 |
| | Pooling_1 | | Kernel size | 2 |
| | | | Padding | Valid |
| | Convolution_2 | Relu | Filters | 128 |
| | | | Kernel size | 5 |
| | Pooling_2 | | Kernel size | 2 |
| | | | Padding | Same |
| | Dropout | | | 0.2 |
| LSTM | LSTM_ | Tach | Neurons | 32 |
| fully connected | Dense_1 | Relu | | 30 |
| | Dense_2 | Softmax | | 7 |
| Training | Optimizer | | | Adam |
| | Batch size | | | 248 |
| | Epochs | | | 80 |

rate of 0.01 to improve the fitting capability, and randomly selected values to initialize the weights and biases of each layer. To increase the robustness model, the order of the training sets was randomly swapped during the training phase. The batch size was set to 80 to increase training effectiveness. The selected hyper-parameters are listed in Table 3.

## Results

The aforementioned CNN–LSTM model was employed for the data collected in this experiment to evaluate the performance of HAR. The data of healthy young group is divided into a training set and test set in a ratio of 7:3. The accuracy of the model is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where $TP$ is the positive samples correctly identified, $TN$ is negative samples correctly identified, $FP$ is the negative samples incorrectly identified and $FN$ is the positive sample incorrectly identified.

Figure 5 shows that with the increase of training epoch, the loss gradually decreases and the accuracy gradually increases, and finally drops to 0.0286 after the training of 80 epoch, whereas the accuracy maximums to 0.9910 after the same training epoch.

Additional metrics like Accuracy, Precision, Recall and F-score were calculated in order to assess the performance of proposed CNN–LSTM model. The definitions of Precision, Recall and F-score are as the following:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

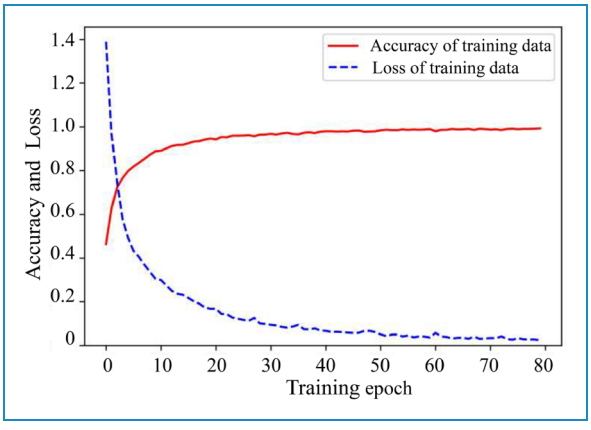


**Figure 5.** The accuracy and loss during the training process on the dataset for the CNN-LSTM model.

$$F - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

A weighted average of Precision and Recall, F-score was introduced since Table 2 shows a class imbalance.

The confusion matrix was used in this study to evaluate the classification results of different models, ablation experiments were added to test the model based on your opinions and the statistical method t-test was used to investigate the variability between the proposed model and the original model. The goal is to demonstrate how each module affects the model; the initial model CNN is supposed to be A, LSTM is B and GRU is C. To study the role of each module, ablation experiments were carried out using A, B, A + B, A + C and A + B + C. The ablation studies yielded good results in CNN–LSTM accuracy, F-score and Param. Table 4 shows the accuracy of each combined model. The A + B + C model is too complex, which is easy to cause overfitting, which affects the classification performance of the model.

An ablation study was necessary for both the healthy person dataset and the pneumoconiosis patient dataset in order to explore the effect of ECG on HAR. We experimented with different input model conditions to see how they affected the classification performance of two separate datasets. To investigate the influence of the ECG, the ACC was fixed and the ECG was raised. In the first step of the ablation experiment, only the influence of ACC on the four depth models was considered. We conducted tests using ACC-only and ACC + ECG combined signal data with four different models (CNN, LSTM, CNN–LSTM and GRU) to assess the impact of ECG inputs on HAR. The confusion matrices of the four models for only ACC signals are shown in Figure 6. The suggested CNN–LSTM model exhibited the best results, as can be observed from the confusion matrix, with an accuracy rate of greater than 0.90 for each behavioral activity. Due to the stark contrast between the activity of lying and other activities detected by accelerometers, there were no false positives

**Table 4.** Accuracy, F-score and Param for different models.

| Model | Accuracy | F-score | Param |
|---|---|---|---|
| A(CNN) | 0.93 | 0.92 | $5.3 \times 10^5$ |
| B(LSTM) | 0.88 | 0.82 | $4.6 \times 10^5$ |
| C(GRU) | 0.94 | 0.94 | $4.2 \times 10^5$ |
| A + B(CNN-LSTM) | 0.98 | 0.98 | $1.7 \times 10^5$ |
| A + C(CNN-GRU) | 0.95 | 0.94 | $2.5 \times 10^5$ |
| A + B + C(CNN-LSTM-GRU) | 0.87 | 0.88 | $6.7 \times 10^5$ |

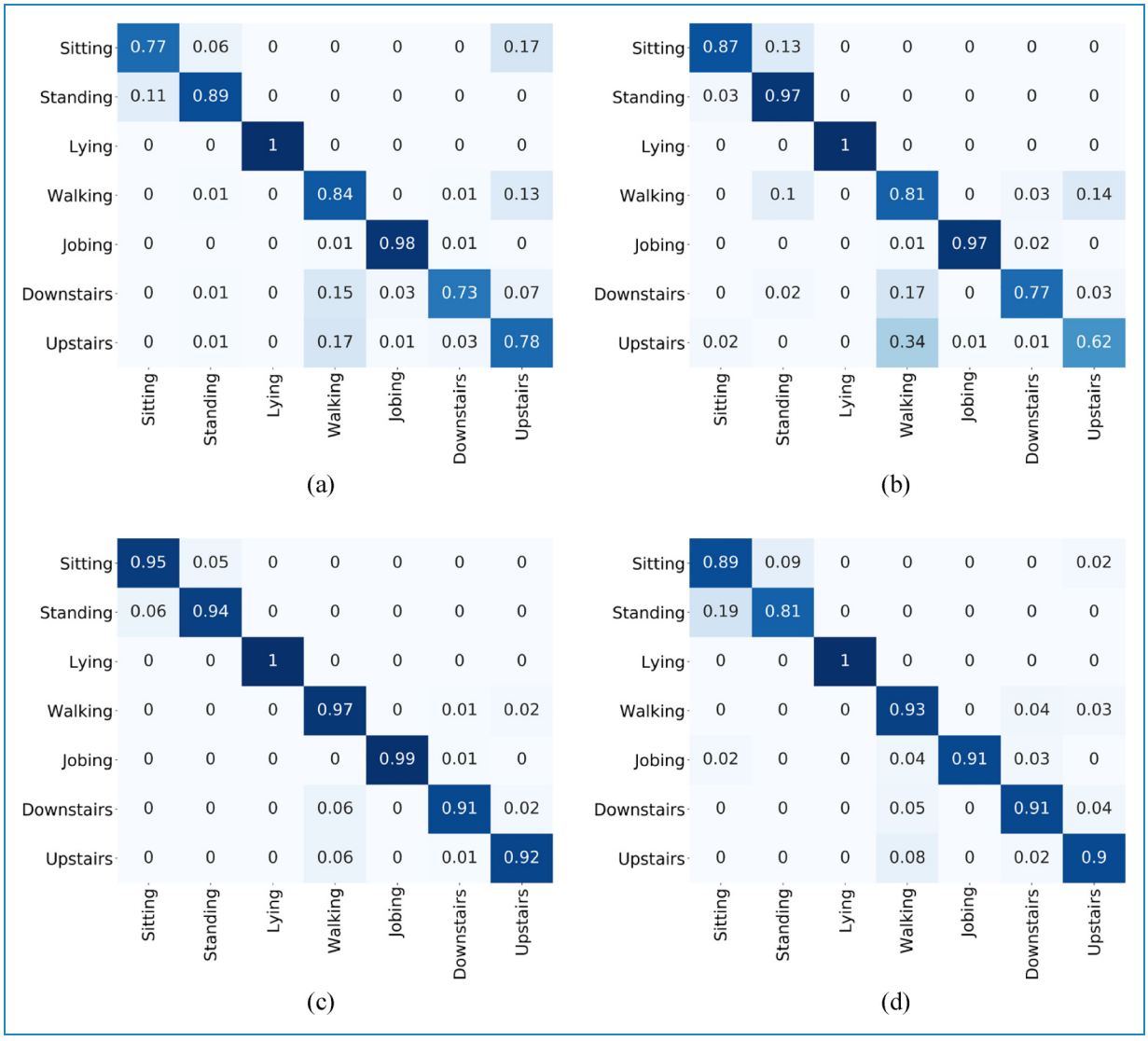


**Figure 6.** Confusion matrix for different models of ACC signals on the dataset of the healthy young group: (a) CNN; (b) LSTM; (c) CNN–LSTM; (d) GRU.

for the lying and its accuracy was 1. It is also noticeable that, when sole ACC was adopted as an input feature, the activities of climbing and descending stairs has slightly worse accuracy than other activities. Dynamic behaviors outperform static behaviors in terms of classification for the whole activities.

In the second half of the ablation study, the ACC input signal is fixed and the ECG signal is added to the ACC signal to examine the effect of ECG on the four models. The confusion matrix for the four models is presented in Figure 7 after combining ACC and ECG inputs to the model. The results also indicate that the CNN–LSTM model outperforms the other three models, achieving an accuracy exceeding 0.97 for each activity. According to the findings, after the addition of the ECG signal, the recognition accuracy of "sitting" in CNN, CNN–LSTM and GRU models grew by 0.18, 0.04 and 0.01, respectively, while the recognition accuracy of "standing" in CNN, LSTM, CNN–LSTM and GRU models increased by 0.04, 0.01, 0.04 and 0.11, respectively.

The comparisons of the overall accuracy and F-score between the only ACC signals and the addition of ECG signals to ACC signals of the four models are displayed in Figure 8. The experimental results demonstrate that incorporating the ECG signal enhances the overall accuracy of the CNN model, LSTM model, CNN–LSTM model and GRU model by 0.0814, 0.0244, 0.0300 and 0.0415, respectively, indicating that the ECG signal has a good impact on the enhancement of HAR accuracy. Meanwhile, the F-score of the dataset obtained from the three models (CNN, CNN–LSTM and GRU) has also increased after combining ECG signals with ACC. with that obtained from CNN–LSTM the

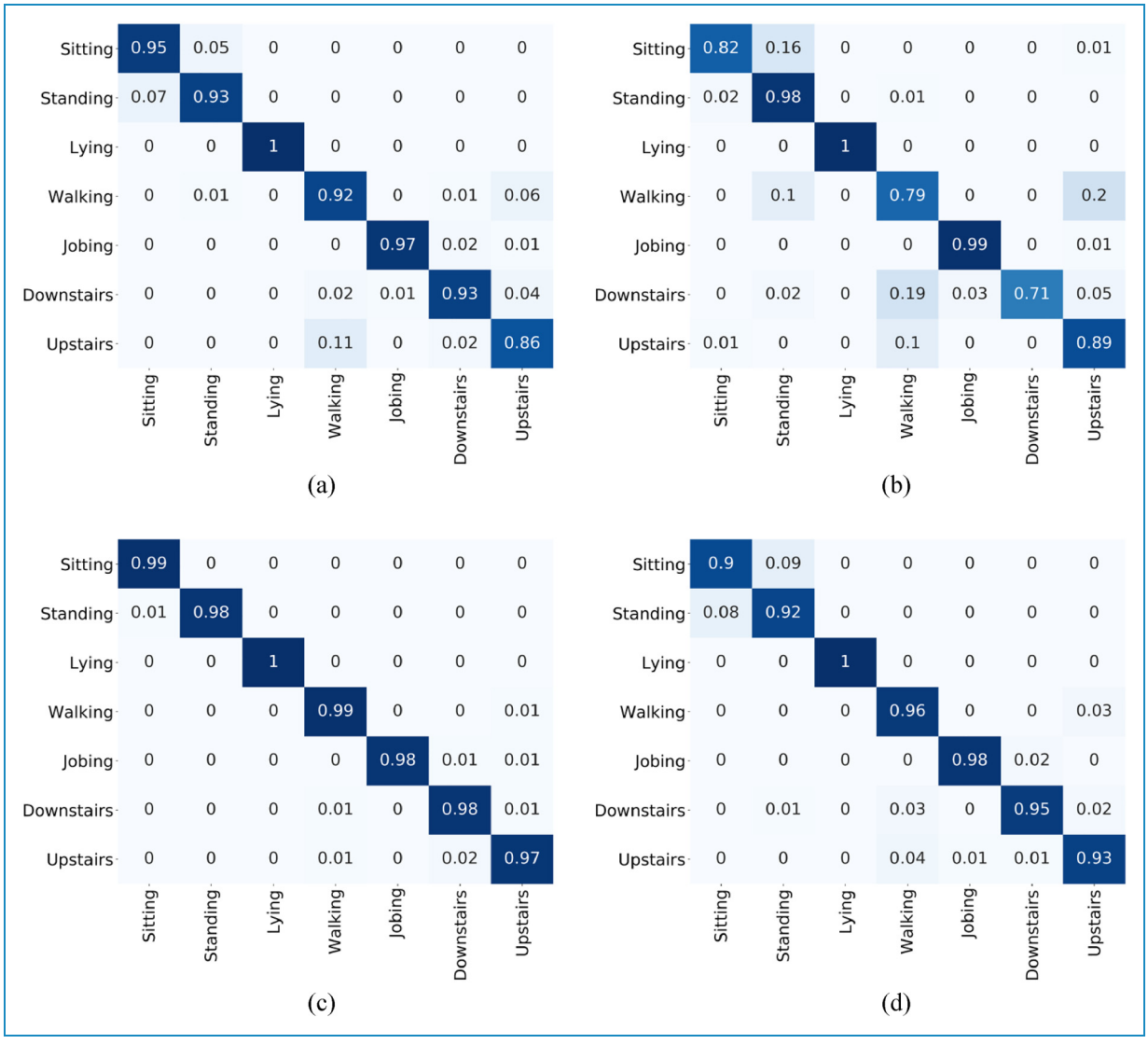


**Figure 7.** Confusion matrix for different models of ACC + ECG signals on the dataset of the healthy young group: (a) CNN; (b) LSTM; (c) CNN-LSTM; (d) GRU.

highest, hitting 0.9840, indicating that the model has best overall classification performance.

Based on the aforementioned ablation experiments, we conducted thorough statistical analyses using SPSS statistical software to discern significant differences among the various models. The t-test was applied to compare different groups, and a two-sided $p$-value of <0.05 was considered indicative of statistically significant differences. The statistical findings revealed noteworthy distinctions between the models ($p < 0.05$), leading to the following conclusions. The amalgamation of CNN and LSTM models demonstrated superior performance compared to utilizing either the CNN or LSTM models in isolation. Notably, when the LSTM model was substituted with the GRU model, the results from the CNN–GRU model did not match the efficacy observed in the CNN–LSTM combination. This underscores the irreplaceable role played by the LSTM model in capturing time series information. In summary, the statistical evidence supports the assertion that the proposed CNN–LSTM model outperforms alternative configurations, emphasizing the significance of integrating convolutional and LSTM architectures for enhanced model performance.

The pneumoconiosis patient dataset also required the same ablation procedures to investigate the effect of ECG addition on HAR. The dataset of the pneumoconiosis patient group was also processed for HAR with each model, and Figure 9 displays the confusion matrix of ACC signals on the dataset of the pneumoconiosis patient group. The activity of "walking" has the highest accuracy for all the four models, reaching 0.86, 0.88, 0.94 and

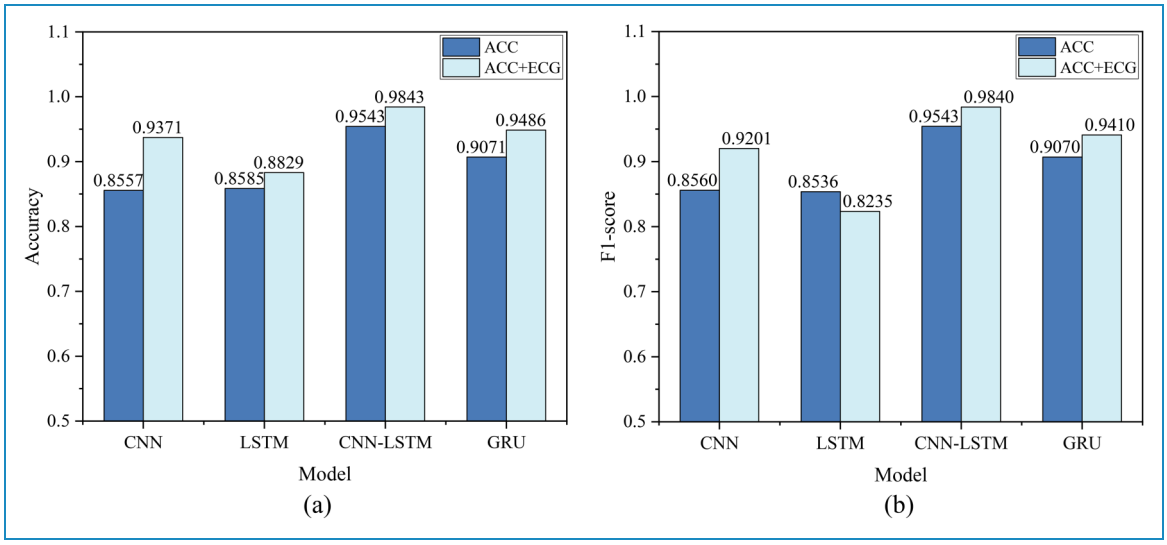


**Figure 8.** Comparison of the (a) accuracy and (b) F-score of different models on the datasets of the healthy young group.

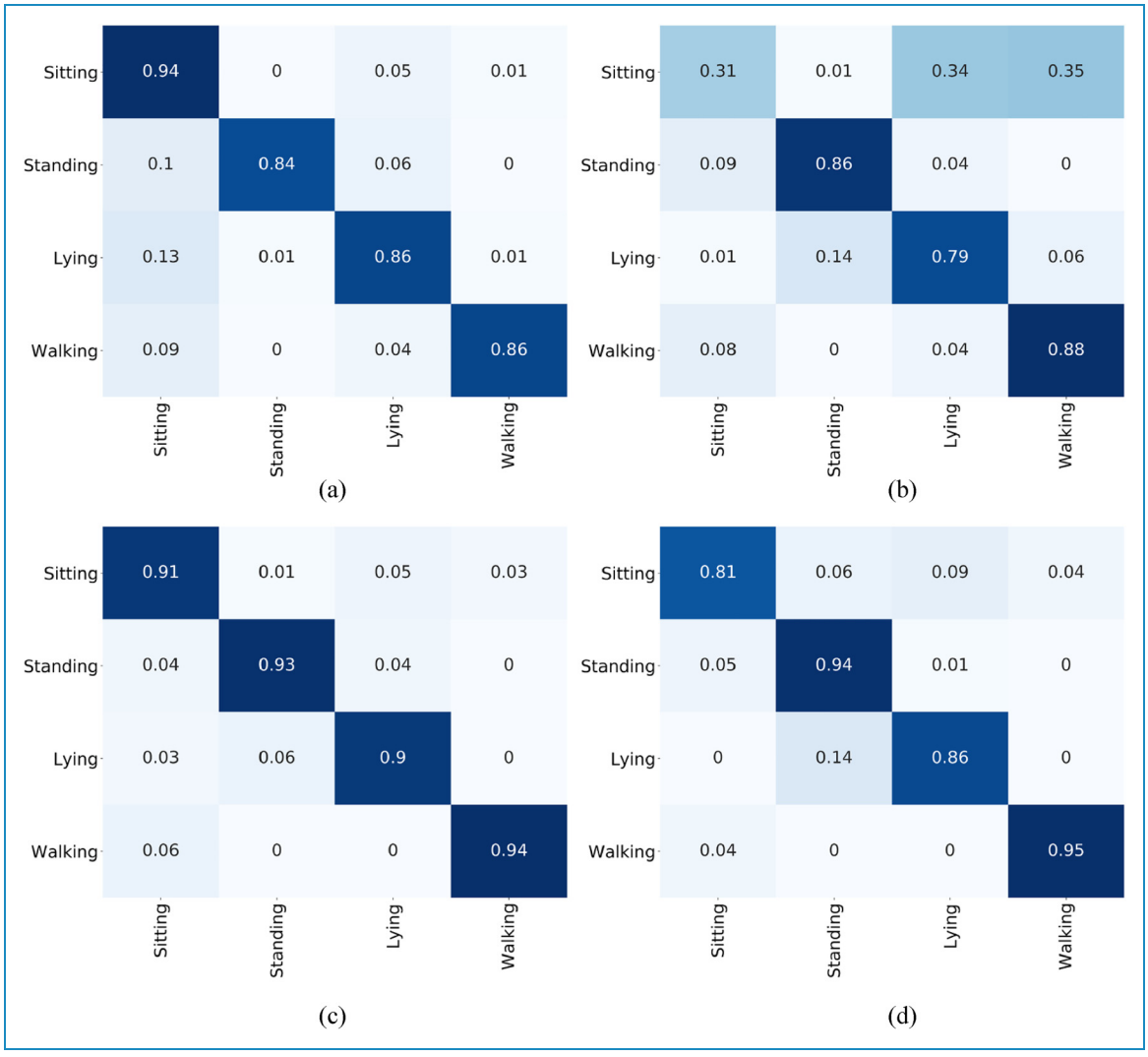


**Figure 9.** Confusion matrix for different models of ACC signals on the dataset of the pneumoconiosis patient group: (a) CNN; (b) LSTM; (c) CNN-LSTM; (d) GRU.


**Figure 10.** Confusion matrix for different models of ACC + ECG signals on the dataset of the pneumoconiosis patient group: (a) CNN; (b) LSTM; (c) CNN–LSTM; (d) GRU.

0.95, respectively, of which the result of the CNN–LSTM model has the highest average recognition accuracy for the four activities, reaching 0.9200.

Figure 10 depicts the confusion matrix that was obtained after analyzing the combination of the ECG and ACC signals of the pneumoconiosis patient group. The results also show that the CNN–LSTM model performs better than other models. The activities of standing and sitting are similar behaviors that are difficult to distinguish, hence sitting has the least accuracy in the four models. When the ECG signal was introduced, the classification accuracy of the "standing" in all four models was improved by 0.04, 0.01, 0.02 and 0.08 percentage points, respectively. However, adding ECG had little effect on "sitting," which has the worst accuracy with only ACC signals as input.

In the pneumoconiosis patient dataset, except for the GRU model, the HAR accuracy and F-score of the other three models (CNN, LSTM, CNN–LSTM) with ECG signals added were higher those with sole ACC signals, as shown in Figure 11. In comparison with the healthy young group, the accuracy and F-score had relatively smaller improvement for the proposed CNN–LSTM model in the pneumoconiosis patient group, with ECG signals added. According to the comparison between Figures 9 and 11, the dataset of the pneumoconiosis patient group had worse overall accuracy than that of healthy young volunteers.

The statistical approach t-test was used to examine the different models, which was consistent with the prior method, and there were also significant differences between the models ($p < 0.05$). Combining CNN and LSTM yields the best classification results because LSTMs can collect time information in the signal and capture longer dependencies than GRU.

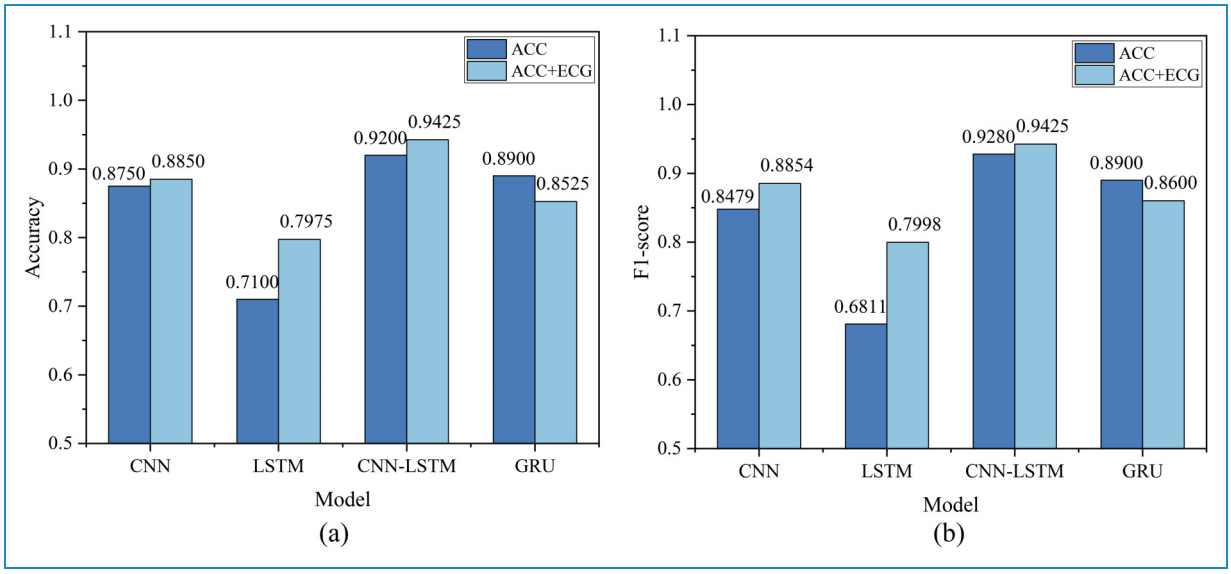


**Figure 11.** Comparison of the (a) accuracy and (b) F-score of different models on the datasets of the pneumoconiosis patient group.

**Table 5.** The model tests the accuracy and F-score of the UCI dataset.

| Literature | Model | Accuracy | F-score |
|---|---|---|---|
| [41] | iSPLInception | 0.95 | 0.95 |
| [42] | DeepConvLSTM | 0.98 | 0.97 |
| [43] | LSTM-CNN | 0.95 | 0.95 |
| Our | CNN-LSTM | 0.99 | 0.99 |

**Table 6.** The model tests the accuracy and F-score of the PAMAP2 dataset.

| Literature | Model | Accuracy | F-score |
|---|---|---|---|
| [41] | iSPLInception | 0.89 | 0.87 |
| [42] | DeepConvLSTM | 0.87 | 0.87 |
| [43] | LSTM-CNN | 0.85 | 0.85 |
| Our | CNN-LSTM | 0.94 | 0.95 |

## Discussion

This study investigated the combination of ACC and ECG signals on clinical HAR, through a wearable patch. The influence of adding ECG on HAR was shown through processing of four deep learning models, especially, a self-constructed CNN–LSTM model was introduced to optimize HAR. The results showed that the CNN–LSTM model achieved the best classification performance in the dataset of the healthy young group and the dataset of the pneumoconiosis patient group. The overall accuracy of HAR in the healthy young group was higher than that in the pneumoconiosis group. In the course of clinical data collection, we discovered that most pneumoconiosis patients have persistent trembling and slow pace due to their ages and medical conditions which caused small fluctuation or noise adding on normal ACC and ECG signals. Therefore, the possibility of misinterpretation will increase in the procedure of data processing by the deep learning models, particularly, in classifying static activities on the dataset. This resulted a lower classification effect in the pneumoconiosis patient group than the healthy young group.

The results demonstrated that ECG signals have a favorable influence on HAR effects for both healthy young group and pneumoconiosis patient group. The HAR accuracy was improved by 0.0814, 0.0244, 0.0300 and 0.0415 point of percentage for the four deep learning models after incorporating ECG signals into ACC signals in the healthy young group, and increased by 0.04, 0.01, 0.02 and 0.08 point of percentage in the pneumoconiosis patient group. However, the degree of improvement in pneumoconiosis patients is not as good as that in healthy people. This also could be caused by the persistent trembling and slow pace during the data collection, resulting in the pneumoconiosis patient group. Since the ECG signal is a faint physiological signal, small fluctuation or noise will largely disturb the ECG signal, and making classification more difficult. Therefore, after adding ECG signals, the improvements of the classification effects for the four models in the

pneumoconiosis patient group were lower than those in the healthy young group.

The relative lower accuracy of HAR and smaller improvement of adding ECG signals in the pneumoconiosis patient group imply that when using HAR models in clinical practice, particular instances need to be further taken into account and optimized. In terms of practical applications, we advise additional study and modification for unique situations like persistent bodily trembling throughout the model training and evaluation phase. Furthermore, additional sensors such as gyroscope sensors to determine how much the body is shaking,[40] are needed in order to further increase the HAR accuracy through deep learning models.

In comparison to the existing dataset that combines acceleration and ECG signals, the data collected in this article reduce the error caused by sensor fusion and increase the activity of lying down, because the activity of lying down is often overlooked, but it plays an important role in the home monitoring of the elderly. It is possible to improve accuracy. Simultaneously, clinical pneumoconiosis patients' data were included to discuss the clinical application, which made a promise for future application prospects. In order to explore the generalization ability of the model proposed in this article, the tests of the public datasets UCI and PAMAP2 were added, and the results obtained were in the Table 5 and Table 6, respectively.

Table 5 shows that the CNN–LSTM model described in this research has the best effect in the UCI dataset, with an accuracy of 0.99 and an F-score of 0.99. Among them, the LSTM–CNN model in Xia et al.[43] feeds the data to the LSTM model for feature extraction, and then to the CNN model. The two-layer LSTM model will increase the amount of calculation, affecting the speed. The model described in this research is first input into the CNN for the building of the spatial information feature map, and then the LSTM model is further refined, which not only improves accuracy but also boosts computing speed.

As can be seen from Table 6, the CNN–LSTM model proposed in this article has the best effect in the PAMAP2 dataset, with an accuracy of 0.94 and an F-score of 0.95.

The comparison results show that the proposed CNN–LSTM model has better accuracy and generalization in HAR.

## Conclusion

This study examines the impact of physiological signal ECG on HAR, and a CNN–LSTM model-based approach is suggested to detect human behavioral activities by examining the dataset gathered by flexible sensors. We tested four depth models in HAR, and the CNN–LSTM provided in the results table performed the best. According to the experimental findings, utilizing both the ECG signal and the ACC signal input model improves the four models' accuracy when compared to using just the ACC signal alone. The CNN–LSTM model excels in classification performance in particular. The healthy volunteer dataset in HAR responded favorably to ECG signals, which might enhance the model's classification performance and increase accuracy by 0.04. Despite the relatively minor influence on the patient data set for pneumoconiosis, the improvement in recognition accuracy is still more objective than the situation without ECG signaling. Due to the significant application potential of ECG as physiological signal in clinical practice, this work argues that the combination of ACC and ECG data is a new HAR technique. The combination of ACC and ECG data has the potential to be useful in future rehabilitation and health management. Medical experts can keep track of a patient's physical activity level, heart health and overall physiological condition to guarantee a smooth recovery while also providing preventative health recommendations. Finally, the integration of ACC and ECG data expands the possibilities for clinical applications, which can increase the accuracy of diagnosis and therapy while also enhancing patient quality of life. This new HAR technology is expected to lead to more medical advancements, offering patients with improved medical treatment and monitoring.

**Acknowledgements**: We are grateful to all who participated in this study.

**Contributorship:** KC and LM got the original ideas and designed the study. YR and ML performed the experiments. YR, ML, YY and KC ran the statistics. YR and KC drafted the manuscript. KC and LM supervised the study.

**Data availability statement:** The data presented in the study are deposited in the Figshare website repository, accessible with the following link: https://doi.org/10.6084/m9.figshare.24132510.v3

**Declaration of conflicting interests:** The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethics approval:** All subjects provided informed written consent prior to participation in the study. This study was conducted in accordance with the guidelines of the World Medical Association Declaration of Helsinki (2000) and was approved and supervised by the Ethics Committee of Shanghai Pulmonary Hospital (Approval Number: L23-226). Written informed consent was obtained from all participants.

**Funding:** The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Key Scientific Research Project of Zhejiang Province, China (Grant No. 2017C03040), Zhejiang Provincial Natural Science Foundation of China (Grant

No. LQ23F030011) and the Fundamental Research Funds for the Provincial Universities of Zhejiang (Grant No. GK22990 9299001-020).

**Guarantor:** KC

**ORCID iDs:** Yanling Ren https://orcid.org/0009-0000-3538-9907
Kai Chen https://orcid.org/0009-0008-6837-3774

## References

1. Bulling A, Blanke U and Schiele B. A tutorial on human activity recognition using body-worn inertial sensors. *Acm Comput Surv* 2014; 46: 1–33.
2. Shavit Y and Klein I. Boosting inertial-based human activity recognition with transformers. *IEEE Access* 2021; 9: 53540–53547.
3. Ghobadi M and Esfahani ET. Foot-mounted inertial measurement unit for activity classification. *IEEE Eng Med Biol Soc* 2014; 2014: 6294–6297.
4. Massé F, Gonzenbach RR, Arami A, et al. Improving activity recognition using a wearable barometric pressure sensor in mobility-impaired stroke patients. *J NeuroEngineering Rehabil* 2015; 12: 72.
5. Lin F, Wang A, Cavuoto L, et al. Toward unobtrusive patient handling activity recognition for injury reduction among at-risk caregivers. *IEEE J Biomed Health Inf* 2017; 21: 682–695.
6. Deepu CJ, Zhang X and Heng CH. A 3-lead ECG-on-chip with QRS detection and lossless compression for wireless sensors. *IEEE Trans Circuits-II* 2016; 63: 1151–1155.
7. Rashkovska A, Depolli M, Tomašić I, et al. Medical-grade ECG sensor for long-term monitoring. *Sensors* 2020; 20: 1695.
8. Chu M, Nguyen T, Pandey V, et al. Respiration rate and volume measurements using wearable strain sensors. *NPJ Digital Med* 2019; 2: 8.
9. Aktaruzzaman M, Scarabottolo N and Sassi R. Parametric estimation of sample entropy for physical activity recognition. *IEEE Eng Med Biol Soc* 2015; 2015: 470–473.
10. Thakur D and Biswas S. An integration of feature extraction and guided regularized random forest feature selection for smartphone based human activity recognition. *J Netw Comput Appl* 2022; 204: 103417.
11. Gomes E, Bertini L, Campos WR, et al. Machine learning algorithms for activity-intensity recognition using accelerometer data. *Sensors* 2021; 21: 1214.
12. Xiao CJ, Lei Y, Ma YS, et al. Deepseg: deep-learning-based activity segmentation framework for activity recognition using WiFi. *IEEE Internet Things* 2021; 8: 5669–5681.
13. Civitarese G, Sztyler T, Riboni D, et al. POLARIS: probabilistic and ontological activity recognition in smart-homes. *IEEE Trans Knowl Data Eng* 2021; 33: 209–223.
14. Lemieux N and Noumeir R. A hierarchical learning approach for human action recognition. *Sensors* 2020; 20: 4946.
15. Murad A and Pyun J-Y. Deep recurrent neural networks for human activity recognition. *Sensors* 2017; 17: 2556.
16. Zheng G. A novel attention-based convolution neural network for human activity recognition. *IEEE Sensors J* 2021; 21: 27015–27025.
17. Wang K, He J and Zhang L. Attention-based convolutional neural network for weakly labeled human activities recognition with wearable sensors. *IEEE Sensors J* 2019; 19: 7598–7604.
18. Lu L, Zhang C, Cao K, et al. A multichannel CNN-GRU model for human activity recognition. *IEEE Access* 2022; 10: 66797–66810.
19. Mekruksavanich S and Jitpattanakul A. RNN-based deep learning for physical activity recognition using smartwatch sensors: a case study of simple and complex activity recognition. *Math Biosci Eng* 2022; 19: 5671–5698.
20. Kaya Y and Topuz EK. Human activity recognition from multiple sensors data using deep CNNs. *Multimed Tools Appl* 2023; 16: 1–19.
21. Li Y, Yang GC, Su ZD, et al. Human activity recognition based on multienvironment sensor data. *Inf Fusion* 2023; 91: 47–63.
22. Soni V, Yadav H, Semwal VB, et al. A novel smartphone-based human activity recognition using deep learning in health care. *Netwk Secur Data Sci* 2023; 946: 493–503.
23. Huang YW, Yang GP, Wang KK, et al. Multi-view discriminant analysis with sample diversity for ECG biometric recognition. *Pattern Recogn Lett* 2021; 145: 110–117.
24. Ma Y, Liu Q and Yang L. Exploring seafarers' workload recognition model with EEG, ECG and task scenarios' complexity: a bridge simulation study. *J Mar Sci Eng* 2022; 10: 1438.
25. Lin Z, Gao Y, Chen Y, et al. Automated detection of myocardial infarction using robust features extracted from 12-lead ECG. *Signal Image Video Process* 2020; 14: 857–865.
26. Kim MG, Ko H and Pan SB. A study on user recognition using 2D ECG based on ensemble of deep convolutional neural networks. *J Ambient Intell Humaniz Comput* 2020; 11: 1859–1867.
27. Salehzadeh A, Calitz AP and Greyling J. Human activity recognition using deep electroencephalography learning. *Biomed Signal Process Control* 2020; 62: 102094.
28. Pawar T, Chaudhuri S and Duttagupta SP. Body movement activity recognition for ambulatory cardiac monitoring. *IEEE Trans Bio Eng* 2007; 54: 874–882.
29. Wahabi S, Pouryayevali S, Hari S, et al. On evaluating ECG biometric systems: session-dependence and body posture. *IEEE Trans Inf Forensics Secur* 2014; 9: 2002–2013.
30. Zou XX, Zhang B and Wang HJ. Study on the current situation of social security for pneumoconiosis in China. *Chin J Ind Hyg Occup Dis* 2021; 12: 954–956.
31. Sun B, Zhao H and Xie Y. Progress in epidemiological studies on pneumoconiosis with comorbidities. *Chin J Ind Hyg Occup Dis* 2021; 39: 389–393.
32. Tseng KK, Fu LA, Liu LL, et al. Human identification with electrocardiogram. *Enterp Inf Syst-Uk* 2018; 12: 798–819.
33. Banos O, Garcia R and Saez A. MHEALTH dataset. UCI machine learning repository. 2014.
34. Merdjanovska E and Rashkovska A. Comprehensive survey of computational ECG analysis: databases, methods and applications. *Expert Syst Appl* 2022; 203: 117206.

35. Cheng J, Chen X and Shen M. A framework for daily activity monitoring and fall detection based on surface electromyography and accelerometer signals. *IEEE J Biomed Health Inf* 2013; 17: 38–45.
36. Cavalieri RN and Pedro BF. Determination of maximum noise level in an ECG channel under SURE wavelet filtering for HRV extraction. *Rev mex ing Bioméd* 2020; 41: 66–72.
37. Noor MHM, Salcic Z and Wang KIK. Adaptive sliding window segmentation for physical activity recognition using a single tri-axial accelerometer. *Pervasive Mob Comput* 2017; 38: 41–59.
38. Wang L and Liu R. Human activity recognition based on wearable sensor using hierarchical deep LSTM networks. *Circ Syst Signal Process* 2020; 39: 837–856.
39. Shu X, Zhang L, Sun Y, et al. Host–parasite: graph LSTM-in-LSTM for group activity recognition. *IEEE Trans Neur Net Lear* 2021; 32: 663–674.
40. Ning Y, Hu S, Nie X, et al. Real-time action recognition and fall detection based on smartphone. *IEEE Eng Med Biol Soc* 2018; 2018: 4418–4422.
41. Ronald M, Poulose A and Han DS. iSPLInception: an inception-ResNet deep learning architecture for human activity recognition. *IEEE Access* 2021; 9: 68985–69001.
42. Ordóñez FG and Roggen D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors* 2016; 16: 115.
43. Xia K, Huang J and Wang H. LSTM-CNN architecture for human activity recognition. *IEEE Access* 2020; 8: 56855–56866.