Computers and Electrical Engineering 121 (2025) 109883

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

A residual deep learning network for smartwatch-based user identification using activity patterns in daily living

Sakorn Mekruksavanich [a], Anuchit Jitpattanakul [b,*]

[a] *Department of Computer Engineering, School of Information and Communication Technology, University of Phayao, 19 Moo 2, Maeka, Muang, Phayao, 56000, Thailand*

[b] *Department of Mathematics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, 1518 Pracharat 1 Rd., Wongsawang, Bangsue, Bangkok, 10800, Thailand*

ARTICLE INFO

Dataset link: https://github.com/GeoTecINIT/sp-sw-har-dataset

Keywords:
User identification
Smartwatch security
Activity patterns
Deep learning
Residual network
Attention mechanism

ABSTRACT

User identification is a critical aspect of smartwatch security, ensuring that only authorized individuals gain access to sensitive information stored on the device. Conventional methods like passwords and biometrics have limitations, such as the risk of forgetting passwords or the potential for biometric data to be compromised. This research proposes a novel approach for user identification on smartwatches by analyzing activity patterns using a hybrid residual neural network called Att-ResBiLSTM. The proposed method leverages unique patterns of user interactions with their smartwatches, including application usage, typing behavior, and motion sensor data, to create an individualized user profile. Employing a deep learning network specifically designed for wearable devices, the system can reliably and promptly identify users by analyzing their activity patterns. The Att-ResBiLSTM architecture comprises three key components: convolutional layers, ResBiLSTM, and an attention layer. The convolutional layers extract spatial features from the pre-processed data. At the same time, the ResBiLSTM component captures long-term dependencies in the time-series data by combining the advantages of bidirectional long short-term memory (BiLSTM) and residual connections. The attention mechanism enhances the final recognition features by selectively prioritizing the most informative elements of the input data. The Att-ResBiLSTM model is trained and evaluated using a diverse dataset of user activity patterns. Experimental results demonstrate that the proposed approach achieves remarkable accuracy in user identification, with an accuracy rate of 98.29% and the highest F1-score of 98.24%. The research also conducts a comparative analysis to assess the efficacy of accelerometer data versus gyroscope data, revealing that combining both sensor modalities improves user identification performance. The proposed methodology provides a reliable and user-friendly alternative to conventional user authentication techniques for smartwatches. This approach leverages activity patterns and a hybrid residual deep learning network to offer a robust and efficient solution for user identification based on smartwatch data, thereby enhancing the overall security of wearable devices.

1. Introduction

Smart wearable devices have become integral to individuals' daily routines in contemporary society, providing various functionalities that extend beyond essential communication. These devices have been of notable use in the Internet of Things (IoT) field. In

* Corresponding author.
E-mail address: anuchit.j@sci.kmutnb.ac.th (A. Jitpattanakul).

https://doi.org/10.1016/j.compeleceng.2024.109883
Received 15 April 2024; Received in revised form 1 October 2024; Accepted 12 November 2024

Available online 27 November 2024

the current scenario, smartphones and smartwatches use their functionalities for data sensing, information representation, and edge computing to facilitate the implementation of IoT solutions [1,2].

Smartphones have grown indispensable in personal and professional settings, often used to access security systems reliant on cloud-based technologies. Smartwatches are advantageous for implementing reliable identity verification systems, mainly when smartphones are susceptible to theft or damage. The use of cloud-based services, such as online banking, may facilitate the achievement of this objective. Ensuring accurate identification of the authentic person responsible for regularly engaging in mission-critical Internet services via cloud-based or alternative data sources is paramount. It is crucial to implement infallible automatic identification systems [3].

Conventional user authentication methods like passwords and biometric identifiers come with inherent flaws [4]. Users may forget their passwords, while biometric data can be vulnerable to unauthorized access. Moreover, some current identification techniques necessitate specific user actions, which can be highly inconvenient and may not accurately capture natural behavior, thereby emphasizing the pressing need for more user-friendly approaches.

Previous studies on wearable device-based user identification have primarily focused on gait analysis [5,6] and touch-based verification [7,8]. However, these approaches, despite their contributions, have their limitations, underscoring the urgent need for more effective methods. Wearable gadgets often struggle with data collection for gait recognition systems due to processing power and battery constraints. Touch-based authentication, on the other hand, relies on specific movements or contact patterns that users might find unintuitive or impractical, further highlighting the need for more user-friendly approaches.

Additionally, existing deep learning-based user identification methods still need to fully exploit the potential of combining multiple sensor inputs. They also need to catch up in capturing long-term patterns in time-series information. For instance, IDNet [9] utilizes a convolutional neural network (CNN) for feature extraction but needs to process one-dimensional sensor data effectively. Giorgi et al.'s research [10] employs a recurrent neural network (RNN) for data classification. However, their findings reveal significant performance variations when differentiating between known and unknown identities, thereby underlining the necessity for more reliable approaches.

In order to overcome these constraints, we provide a novel method for user authentication on smartwatches. This method examines activity patterns using a hybrid residual neural network called Att-ResBiLSTM. The suggested method utilizes distinct user engagement patterns with their smartwatches to establish personalized user profiles. The system utilizes a specialized deep learning network specifically for wearable electronics to quickly and accurately identify users based on their activity patterns.

This research proposes using a hybrid residual deep learning network named Att-ResBiLSTM to enhance identification accuracy. The proposed network architecture integrates convolutional layers, a ResBiLSTM component, and an attention mechanism. The amalgamation of these constituents amplifies the network's capacity to extract spatial characteristics, capture enduring relationships in time series data, and selectively concentrate on the most informative segments of the input data, thereby enhancing accuracy and consistency of user identification derived from smartwatch sensor data. An assessment was conducted on the following areas and contributions:

- The Att-ResBiLSTM model was proposed as an innovative hybrid residual network that combines convolutional layers to extract spatial features, the ResBiLSTM component to capture long-term relationships in time series data, and an attention mechanism to optimize the final recognition features.
- An investigative study was conducted to evaluate the effectiveness of accelerometer data against gyroscope data. The findings revealed that amalgamating both sensor types improves user identification outcomes.
- Our Att-ResBiLSTM model surpassed other deep learning methodologies on the SP-SW HAR dataset. It achieved a remarkable accuracy rate of 98.29% and attained the highest F1-score of 98.24%.

This paper is structured as follows: Section 2 offers an extensive survey of prior research on machine learning and deep learning techniques for examining mobile sensor data and user identification. Section 3 outlines the framework for biometric user recognition utilizing wearable device sensing. Section 4 elucidates the experimental setup, while Section 5 evaluates and discusses the results obtained. Lastly, Section 6 presents the study's conclusions, limitations, and potential avenues for future exploration.

2. Related works

This part provides a concise overview of the study on sensor-based user identification. The following sections provide further information.

2.1. Sensor-based user identification

The last decade has witnessed a surge in the adoption of wearable sensor-driven monitoring systems that rely on smartwatches. One such instance is the system concept proposed by Saini et al. [11], designed to openly and consistently verify an individual's identity. However, in most cases, a dearth of evidence supports the claim that an individual's behavior has undergone alterations substantial enough to justify its unambiguous classification. Luca et al. [12] introduced a methodology that intentionally quantifies the divergence between pattern traces by utilizing the dynamic temporal warping mechanism. The research conducted by Sae-Bae et al. [13] demonstrated that most common touch patterns necessitate the simultaneous engagement of all five fingers. The study in [14] employed classification algorithms such as support vector machines and k-nearest neighbors to detect the analytical features extracted from touch recordings.

Table 1
A summary of existing literature on user identification based on the sensor data.

Work (Year)	Classifier	Sensors	Device	Performance (% accuracy)	No. of users
Mekruksavanich et al. [3] (2022)	1D-ResNet-SE	1 Acc. 1 Gyro.	Smartwatch	95.37	51
Benegui et al. [27] (2020)	CNN	1 Acc. 1 Gyro.	Smartphone	90.75	50
Angrisano et al. [28] (2020)	Random Forest	1 Acc. 1 Gyro.	Smartphone	93.8	32
Weiss et al. [29] (2019)	Random Forest	1 Acc. 1 Gyro.	Smartphone	92.7	51
		1 Acc. 1 Gyro.	Smartwatch	71.7	
Musale et al. [30] (2019)	Random Forest	1 Acc. 1 Gyro.	Smartwatch	91.8	51
Ahmad et al. [31] (2018)	Decision Tree	1 Acc. 1 Gyro. 1 Mag.	Smartwatch	98.68	6
Neverova et al. [32] (2016)	CNN	1 Acc. 1 Gyro. 1 Mag.	Smartphone	69.41	587

Consequently, within the behavioral paradigm, temporal and spatial characteristics are the two primary attributes of each response. For instance, when attempting to recognize a user, one may consider approaches similar to those outlined in [15]. The notion of multi-modal continuous user identification was previously introduced by Sabharwal [16]. An alternative architectural framework for continuous user identification, leveraging historical smartphone data and location information, was proposed in the research work [17].

The previously mentioned tasks necessitate additional data and a user authentication mechanism. To tackle these challenges, Casale et al. [18] proposed a user identification system that leverages an inconspicuous biometric pattern based on an individual's gait.

2.2. Deep learning approaches in user identification

Gait recognition, alongside motion detection [19], video classification [20], and facial recognition [21], is a domain where deep learning has demonstrated its efficacy. However, acquiring gait data can be challenging due to the limited computational power and battery life of portable devices, rendering the training of deep learning models a data-intensive process. The application of deep learning techniques to address gait-based implicit authentication challenges is less prevalent than traditional machine learning approaches. Gadaleta et al. [9] introduced IDNet, an authentication system for users that utilizes smartphone location data. To extract generalizable features, IDNet employs CNN, while the classifier remains a one-class support vector machine. The initial gait signal is directly fed into the CNN. However, this approach must be better-suited for CNN implementation, as CNNs inherently struggle to process one-dimensional data effectively. Giorgi et al. [10] developed a user authentication method that leverages inertial sensors and a RNN for deep learning-based classification. Nevertheless, the results for known identities and unknown identities exhibit significant discrepancies.

In comparison to recent scholarly inquiries, there are notable advantages to be found in cross-channel interaction when it comes to boosting the effectiveness of deep learning methods. A novel approach known as the squeeze-and-excitation (SE) network, introduced by Hu et al. [22], stands out as a means to tweak and refine the responses of channel features. Researchers like Chen et al. [23] and Dai et al. [24] have employed attention techniques at the channel level in their investigations into semantic segmentation and picture captioning. Wu and He [25], on the other hand, utilized a method called group normalization. This particular model can be seen as a tailored solution incorporating channel-based connections. However, the complexity arises in how these channels interact across multiple channels, as it involves more than just computing feature maps' averages and standard deviations. In computer vision, a concept proposed by Yang et al. [26] suggests integrating cross-channel interaction within a single layer. This approach fosters inter-channel communication within the same layer, enhancing overall effectiveness.

2.3. State-of-the-art works

Numerous new methodologies and sensor modalities have recently been proposed for user identification through smartwatches, utilizing state-of-the-art deep learning technologies. Table 1 provides a concise overview of these advanced methods, emphasizing their main characteristics, structures, and effectiveness measurements.

Several recent studies have concentrated on user identification through sensor data analysis, each taking a unique approach. Mekruksavanich et al. [3] employed a 1D-ResNet-SE classifier on a smartwatch accelerometer and gyroscope data to identify users

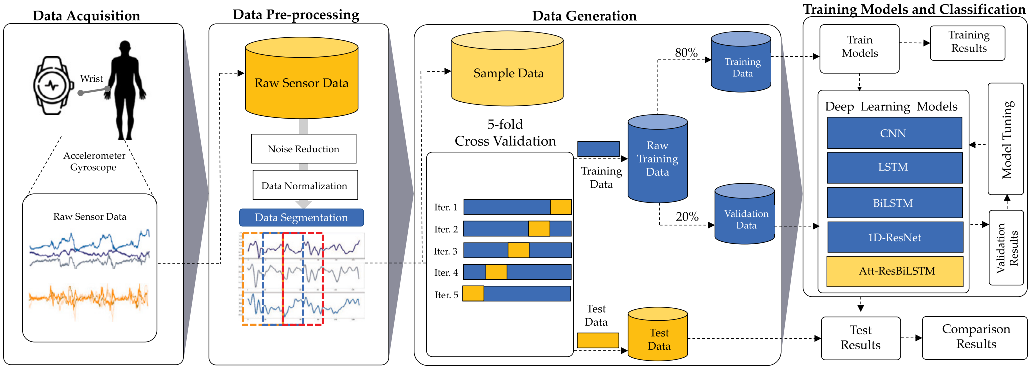


Fig. 1. The proposed framework of activity-based user identification using smartwatch sensors.

during dynamic activities, achieving 95.37% accuracy with 51 participants. Similarly, Benegui et al. [27] developed a system using smartphone motion sensor data from screen-tapping behaviors, utilizing a CNN classifier and reaching 90.75% accuracy with 50 users. Focusing on walking habits, Angrisano et al. [28] used a random forest classifier on smartphone accelerometer and gyroscope data, achieving 93.8% accuracy with 32 individuals. They enhanced accuracy through ensemble machine-learning techniques. Musale et al. [30] also analyzed walking patterns using smartwatch sensors, employing a random forest classifier and statistical features derived from sensor data, attaining 91.8% accuracy with 51 individuals. Weiss et al. [29] took a comprehensive approach, exploring user identification through both simple and sophisticated actions. They used a random forest classifier on data from smartphones and smartwatches, achieving 92.7% accuracy and 71.7% for smartwatch data with 51 users. Ahmad et al. [31] developed a system based on ambulatory behaviors using wristwatch sensors, employing a decision tree classifier on accelerometer, gyroscope, and magnetometer data. They achieved a remarkable 98.68% accuracy, albeit with a smaller dataset of 6 users. Lastly, Neverova et al. [32] introduced a technique using deep convolutional clockwork RNNs to analyze walking movements. Using a CNN classifier on smartphone sensor data, they achieved 69.41% accuracy on a large dataset of 587 individuals. These studies not only demonstrate the potential of sensor-based user identification, but also the diversity of approaches across various devices and activities. The accuracies range from 69.41% to 98.68%, highlighting the impact of methodology and dataset size on the results.

Recent studies highlight the effectiveness of machine learning and deep learning methods in identifying users by analyzing sensor data collected from smartphones and smartwatches. These techniques employ various classifiers, such as CNN, long short-term memory (LSTM) networks, random forests, and decision trees, utilizing multiple sensor modalities, including accelerometer, gyroscope, and magnetometer data.

However, challenges and opportunities for further research remain in this domain. Prospective study efforts should focus on enhancing the robustness and adaptability of user identification models, especially for more extensive and diverse user groups. It is also essential to develop more efficient and less resource-intensive models compatible with wearable devices with limited computational power.

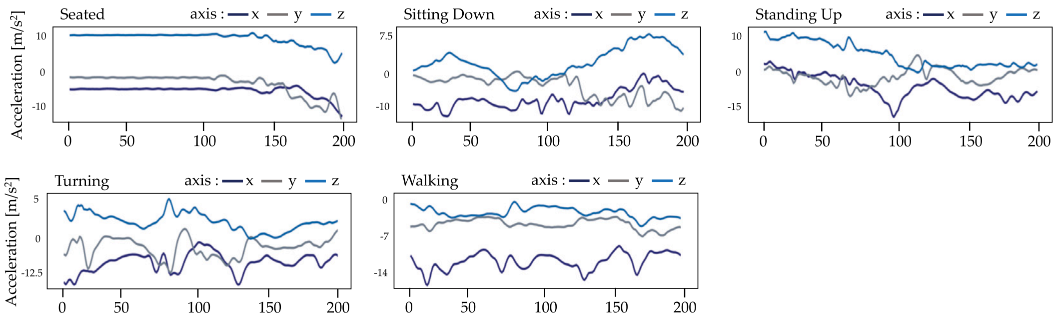
3. Activity-based user identification framework

In this study, the framework employed for user identification relies on activity patterns. This framework encompasses four primary phases: data gathering, pre-processing, generation, and training and evaluation of the model, as depicted in Fig. 1. Each of these stages will be detailed in the subsequent sections.

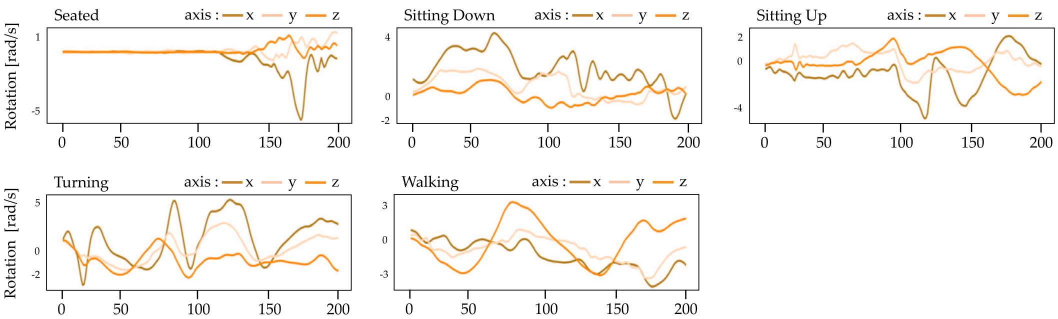
The activity-based user identification framework using smartwatch sensors involves four primary steps: data gathering, data pre-processing, the proposed Att-ResBiLSTM model, and performance evaluation. This framework is not just a theoretical concept, but a practical solution that can be applied in real-world scenarios. In the data acquisition stage, raw sensor data is collected from smartwatches worn by individuals engaged in various everyday activities. During the pre-processing phase, the data is segmented, noise is removed, and the data is standardized to prepare it for training and evaluation. The Att-ResBiLSTM model, a hybrid deep learning network, is designed to classify individuals by analyzing their activity patterns. This model comprises convolutional layers, a ResBiLSTM block, and an attention layer. The model's performance is assessed by measuring parameters such as accuracy, precision, recall, and F1-score using a 5-fold cross-validation procedure on the SP-SW HAR dataset.

3.1. Data acquisition

This study explored user identification using data collected from smartwatch sensors within the SP-SW HAR dataset [33], focusing on five distinct activity patterns. This dataset serves as a widely accessible benchmark for human activity recognition (HAR), containing sensor data from both smartphones and smartwatches, specifically the inertial measurement unit (IMU). The dataset includes records of linear acceleration and angular velocity from a sample of 23 individuals aged 23 to 66, with 10 females and 15 males represented. Participants were equipped with a smartphone and a wristwatch, outfitted with the necessary software, and



(a) Acclerameter data



(b) Gyroscope data

Fig. 2. Some sample data of five activities from the SP-SW HAR dataset.

instructed to carry out their daily activities without constraints. The participants categorized tasks such as sitting, standing, turning, and walking as part of the data collection process. Fig. 2 illustrates various instances of accelerometer and gyroscope data. We chose to prioritize data collected from smartwatch sensors for this study for several reasons. Smartwatches are ubiquitous in modern life, making them an ideal platform for capturing user behavior. Moreover, smartwatches boast enhanced computational capabilities, enabling deploying highly efficient deep learning models directly on the device.

3.2. Pre-processing process

During the data collection process, the participants' active movements introduced various sources of interference and noise to the raw sensor data. Signals containing high noise levels tend to become distorted, making it imperative to minimize noise's impact during signal processing to extract meaningful information [34]. Filtering techniques such as mean, low-pass, and Wavelet filtering are commonly employed. For our research, we used a third-order Butterworth filter with a cutoff frequency of 20 Hz to remove unwanted noise from accelerometers, gyroscopes, and magnetometers in all three dimensions. This filter effectively captures nearly all body motions, accounting for 99.9%, rendering it highly suitable for motion recording applications [35].

After removing unnecessary noise, it was imperative to modify the sensor data. Every data point underwent a Min-Max normalization technique, which maps its values into the interval [0, 1]. Establishing a method to mitigate the effects of various factors might be advantageous for the educational processes. The approach involves segmenting the normalized data from all sensors into equal-sized portions using fixed-size sliding windows for model training. In this investigation, we used a sliding window with a duration of 10 s, as recommended by Weiss wt al. [29], to create sensory data streams of a certain length. The 10-second window is used to identify users because it is enough time to collect crucial attributes of an individual's behaviors, such as recurring fundamental motions like walking and ascending stairs. This permits prompt biometric recognition. Moreover, prior research on activity detection has shown that a window duration of 10 s achieves superior performance compared to other durations [36].

3.3. The proposed Att-ResBiLSTM architecture

Our research introduces a new method called Att-ResBiLSTM, a hybrid residual neural network tailored to address the user identification problem. The model's architecture, illustrated in Fig. 3, comprises three key components: convolutional layers, ResBiLSTM, and an attention layer. The convolutional layer first captures spatial characteristics from the pre-processed input. Efficiently modifying the stride of the convolution kernel decreases the length of the time series, resulting in faster recognition.

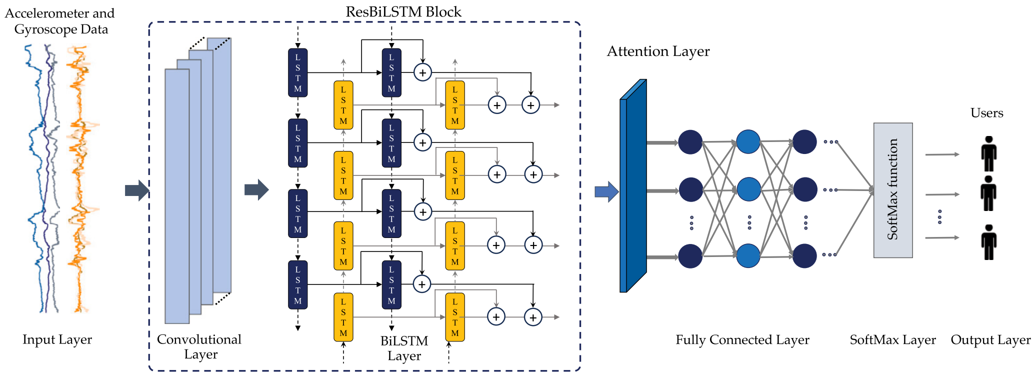


Fig. 3. The architecture of the Att-ResBiLSTM model.

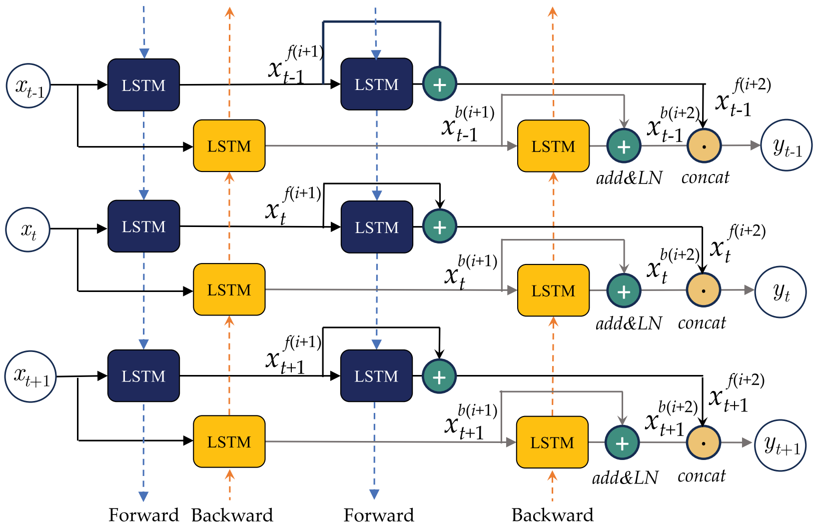


Fig. 4. Structure of the ResBiLSTM.

Subsequently, the improved ResBiLSTM network is used to extract temporal patterns from the input that the convolutional layer has processed. The ResBiLSTM component improves the model's capacity to capture long-term dependencies in time series data by using bidirectional LSTM (BiLSTM) and integrating residual connections. This integration enhances the model's comprehension of intricate temporal patterns and recognition precision. In order to enhance the accuracy of recognition, we suggest including an attention mechanism. This process calculates weights for the feature information produced by the ResBiLSTM network, allowing the model to concentrate on the most informative parts of the input data. The attention mechanism improves user identification accuracy by highlighting essential characteristics and strengthening the model's ability to differentiate across users.

3.3.1. Convolutional layers

The proposed approach employs a one-dimensional CNN (1D-CNN) to extract features from sensor data effectively. The Smish activation function is used to handle the issue of negative interval. At the same time, batch normalization (BN) is applied to reduce the problem of internal covariate shift and accelerate convergence during training. The Att-ResBiLSTM model includes a convolution block consisting of Conv1D, BN, Smish, and max-pooling layers. This convolution block is repeated four times within the model.

3.3.2. Residual BiGRU block

The ResBiLSTM model is designed to process sequential data and detect bidirectional dependencies. This model builds on the bidirectional LSTM (BiLSTM) network by incorporating a residual framework. This strategy addresses the vanishing gradients problem and enhances the feature extraction from time series data. Instead of batch normalization (BN), layer normalization (LN) is used in the ResBiLSTM to boost performance. Fig. 4 provides a detailed view of the ResBiLSTM's architecture.

3.3.3. Attention block

The model includes an attention mechanism to identify spatial relationships within the data and assess the importance of various information components. This mechanism enhances the precision of behavior prediction by modifying the weights throughout the training process. As training progresses, the feature information becomes increasingly representative, leading to improved accuracy in recognition.

3.4. Performance evaluation

Many performance measures, including accuracy, precision, recall, and F1-score, are used to assess the effectiveness of the identification models. The performance metrics used are delineated as mathematical equations below.

$$Precision\ (\%) = \frac{TP}{TP + FP} \times 100\% \tag{1}$$

$$Recall\ (\%) = \frac{TP}{TP + FN} \times 100\% \tag{2}$$

$$F1 - score\ (\%) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% \tag{3}$$

$$Accuracy\ (\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{4}$$

These four evaluation metrics are frequently employed in user identification studies to gauge their effectiveness. A correct identification is labeled as a true positive (TP) for the specific category being assessed and a true negative (TN) for all other categories. Additionally, sensor data from one category may be incorrectly classified as belonging to another, leading to false positives (FP) for that category. Conversely, sensor data from another category may be misclassified as belonging to a different category, resulting in false negatives (FN) for that category's identification.

4. Experiments and results

This section presents the outcomes of all the tests we ran to discover effective deep-learning models for identifying users using smartwatches. The study used the SP-SW HAR dataset to identify individuals by analyzing data collected from smartwatches. The deep learning models were assessed for accuracy, precision, recall, and F1-score using these metrics.

4.1. Environmental configuration

This experiment utilized the Google Colab Pro+ platform to conduct deep learning model training. Accelerated training was achieved by leveraging the Tesla V100-SXM2, which features a 16 GB graphics processing unit, resulting in impressive outcomes. To incorporate the Att-ResBiLSTM and other foundational deep learning models into the Python ecosystem, we opted for the Tensorflow backend 3.9.1 and CUDA 8.0.6 graphics cards. The Python libraries involved in this study encompassed various tasks:

- Numpy and Pandas managed sensor data, including reading, processing, and assessing it.
- Matplotlib and Seaborn were utilized to visualize and present the results of data exploration and model evaluation processes.
- Sklearn, a specialized sampling and data generation library, played a crucial role in the study.
- The development and training of deep learning models included using TensorFlow, Keras, TensorBoard, and additional programming languages.

4.2. Experiment setting

In this study, we conducted five distinct experiments using sensor data from the SP-SW HAR dataset, as detailed in Table 2. The proposed Att-ResBiLSTM model was evaluated for its identification performance across these five trials, employing a 5-fold cross-validation methodology.

In deep learning, hyperparameter variables are crucial in regulating and fine-tuning the learning process. The Att-ResBiLSTM model incorporates several vital hyperparameters, including:

1. The number of epochs,
2. The batch size,
3. The learning rate (denoted as α),
4. The optimization method,
5. The loss function.

Table 2
Experiments cases in this study.

Experiment case	Description
Experiment I	This revolved around training and assessing deep learning models using accelerometer and gyroscope data captured from users' smartwatches during turning movements.
Experiment II	This involved training and evaluating deep learning models using sensor data obtained from smartwatches while users were walking.
Experiment III	This concentrated on training and evaluating deep learning models with sensor data collected from smartwatches worn by participants in a seated position.
Experiment IV	This focused on training and assessing deep learning models with data from smartwatches while users were seated.
Experiment V	This focused on training and assessing deep learning models using sensor data gathered from smartwatches while users remained stationary.

Table 3
Experimental results of deep learning models for turning actions using smartwatch sensor data.

Sensor	Identification performance (%mean(±Std.))			
	Accuracy (↑)	Precision (↑)	Recall (↑)	F1-score (↑)
Acc.	78.65%(±7.48%)	81.87%(±5.99%)	79.01%(±7.70%)	78.31%(±8.14%)
Gryo.	79.88%(±2.99%)	82.04%(±2.63%)	79.13%(±2.88%)	78.98%(±3.03%)
Acc.+Gyro.	88.40%(±0.82%)	89.79%(±0.22%)	88.00%(±0.78%)	87.86%(±0.66%)

Note: Acc. is Accelerometer data and Gryo. is Gyroscope data. (↑) and (↓) indicate higher and lower values are better for these metrics, respectively.

For our experimentation, we configured these hyperparameters as follows: setting the number of epochs to 200 and the batch size to 128. Additionally, to prevent overfitting, we implemented an early-stop callback mechanism, terminating training if validation loss did not decrease after 20 epochs.

Initially, we set the learning rate (α) to 0.001. However, if the validation accuracy did not improve for six consecutive epochs, we reduced it by 25% of its original value.

To optimize the learning process, we employed the Adam optimizer [37] with specific parameter values: $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 1 \times 10^8$. The categorical cross-entropy function was utilized to assess the optimizer's performance. Recent research has demonstrated the superiority of the cross-entropy approach over alternative strategies such as classification error and mean square error [38].

4.3. Experimental findings

In this section, we present the experimental results of our proposed Att-ResBiLSTM model, along with other deep learning models, for user identification based on smartwatch sensor data. Five experiments focused on a specific activity: rotational movements, walking, sitting, standing, and sitting down. These experiments employed a 5-fold cross-validation technique to evaluate the models' identification performance.

The sensor data in these experiments were sourced from the publicly available SP-SW HAR dataset, including accelerometer and gyroscope data captured from users' smartwatches during various activities. The performance of the models was assessed using multiple metrics, including accuracy, precision, recall, and F1-score.

The results indicate that the Att-ResBiLSTM model excels at capturing distinct patterns associated with different activities for user identification. Furthermore, leveraging data from multiple sensors offers additional benefits in enhancing identification accuracy.

The Att-ResBiLSTM model outperformed other deep learning models in utilizing accelerometer and gyroscope data from smartwatches worn during turning activities, as shown in Table 3. The table presents performance metrics such as accuracy, precision, recall, and F1-score for three sensor configurations: accelerometer, gyroscope, and combination.

Analysis reveals that using only accelerometer data resulted in an accuracy range of 78%–79%. Among these models, the Att-ResBiLSTM achieved the highest F1-score of 78.31%. Utilizing gyroscope data alone showed a slight improvement, with an accuracy of approximately 79%–80%. Moreover, the Att-ResBiLSTM model achieved an F1-score of 78.98% in this configuration. The most optimal performance was achieved by integrating both accelerometer and gyroscope data. So, the Att-ResBiLSTM model achieved an accuracy of 88.40% and the highest F1-score of 87.86%, surpassing other models.

These results highlight the advantage of integrating data from multiple sensors, specifically accelerometer and gyroscope, for improved person identification compared to using a single sensor type. The proposed Att-ResBiLSTM model demonstrates superior effectiveness in capturing unique patterns in users' turning movements, resulting in higher levels of accuracy and F1-score compared to other deep learning models.

Table 4 presents the performance of deep learning models, including the proposed Att-ResBiLSTM model, for user identification using smartwatch sensor data during walking activities. The models were evaluated with three sensor configurations: accelerometer alone, gyroscope alone, and a combination of both.

Using only accelerometer data, the models achieved accuracies ranging from 92.24% to 92.55%, with Att-ResBiLSTM having the highest F1-score of 92.25%. With gyroscope data alone, performance improved, with accuracies between 96.30% and 96.33%, and

Table 4
Experimental results of deep learning models for walking actions using smartwatch sensor data.

Sensor	Identification performance (%mean(±Std.))			
	Accuracy (↑)	Precision (↑)	Recall (↑)	F1-score (↑)
Acc.	92.25%(±10.00%)	94.00%(±6.99%)	92.55%(±9.10%)	92.24%(±9.93%)
Gryo.	96.33%(±0.65%)	96.60%(±0.50%)	96.30%(±0.73%)	96.33%(±0.60%)
Acc.+Gyro.	98.29%(±0.73%)	98.44%(±0.53%)	98.19%(±0.85%)	98.24%(±0.74%)

Note: Acc. is Accelerometer data and Gryo. is Gyroscope data. (↑) and (↓) indicate higher and lower values are better for these metrics, respectively.

Table 5
Experimental results of deep learning models for seated activities using smartwatch sensor data.

Sensor	Identification performance (%mean(±Std.))			
	Accuracy (↑)	Precision (↑)	Recall (↑)	F1-score (↑)
Acc.	26.70%(±7.82%)	15.70%(±6.55%)	23.67%(±8.08%)	17.66%(±6.61%)
Gryo.	8.87%(±3.91%)	2.30%(±2.75%)	9.09%(±4.30%)	3.11%(±3.09%)
Acc.+Gyro.	29.97%(±11.12%)	20.13%(±8.20%)	25.58%(±8.82%)	20.76%(±7.95%)

Note: Acc. is Accelerometer data and Gryo. is Gyroscope data. (↑) and (↓) indicate higher and lower values are better for these metrics, respectively.

Table 6
Experimental results of deep learning models for sitting activities using smartwatch sensor data.

Sensor	Identification performance (%mean(±Std.))			
	Accuracy (↑)	Precision (↑)	Recall (↑)	F1-score (↑)
Acc.	77.09%(±5.81%)	78.86%(±3.07%)	74.05%(±4.98%)	72.94%(±4.99%)
Gryo.	33.87%(±4.11%)	35.88%(±3.78%)	32.64%(±2.87%)	26.34%(±3.78%)
Acc.+Gyro.	89.46%(±4.00%)	90.71%(±3.46%)	88.55%(±3.76%)	87.73%(±4.13%)

Note: Acc. is Accelerometer data and Gryo. is Gyroscope data. (↑) and (↓) indicate higher and lower values are better for these metrics, respectively.

Att-ResBiLSTM achieved an F1-score of 96.33%. The best performance was obtained by combining accelerometer and gyroscope data. The Att-ResBiLSTM model outperformed other models, achieving an accuracy of 98.29% and the highest F1-score of 98.24%.

The findings indicate that combining multiple sensor data improves user identification compared to using a single sensor type. The Att-ResBiLSTM model effectively captures distinctive patterns in users' walking behaviors, leading to higher accuracy and F1-scores. User identification based on walking activities performs better than that based on turning movements (Table 3), suggesting that walking patterns are more distinctive and identifiable.

Table 5 presents the performance of deep learning models for user identification based on accelerometer and gyroscope data collected from smartwatches during sitting activities. The results show significantly poorer performance than user identification based on turning (Table 3) and walking (Table 4) activities.

The models' accuracy for sitting activities ranges from 8.87% to 29.97%, with the highest F1-score being just 20.76% when combining accelerometer and gyroscope data.

This unsatisfactory performance suggests that sitting activities must provide distinctive patterns for effective user identification. The limited body movements during sedentary activities make it challenging for the models to capture unique patterns for each user.

These findings highlight the importance of selecting appropriate activities and sensor data for accurate user identification using wearable devices. Sitting activities are unsuitable for this purpose.

Table 6 presents the empirical results of deep learning models utilizing accelerometer and gyroscope sensor data collected from smartwatches worn by users during sedentary periods. The models' performance surpasses the data depicted in Table 5 for sitting activities. However, it still falls short of the outcomes observed for rotational movements in Table 3 and walking activities in Table 4.

The models' accuracy ranges from 33.87% to 89.46%, with the Att-ResBiLSTM model achieving the highest F1-score of 87.73% when amalgamating accelerometer and gyroscope sensors data. These findings suggest that sitting activities generate more distinctive patterns than actions performed in a seated position, enabling more precise user identification. Nonetheless, the performance of sitting activities remains inferior to that of rotational movements and ambulatory activities, implying that there may be more reliable approaches for user identification.

Table 7 presents the observed results of deep learning models utilizing accelerometer and gyroscope data gathered from users' smartwatches while in a stationary position. Although the models' performance resembles that of sitting activities (Table 6), it falls short of the outcomes for turning movements (Table 3) and walking (Table 4).

Model accuracy ranges from 34.60% to 91.61%, with the Att-ResBiLSTM model achieving the highest F1-score of 90.64% when utilizing both accelerometer and gyroscope data. This indicates that stationary activities yield somewhat distinct patterns compared to sitting, but they may not be as reliable as turning or walking activities for user identification.

Table 7
Experimental results of deep learning models for standing activities using smartwatch sensor data.

Sensor	Identification performance (%mean(±Std.))			
	Accuracy (↑)	Precision (↑)	Recall (↑)	F1-score (↑)
Acc.	84.29%(±1.92%)	87.66%(±1.87%)	83.97%(±2.08%)	82.96%(±1.92%)
Gryo.	34.60%(±5.05%)	35.82%(±4.53%)	30.93%(±3.79%)	27.02%(±3.83%)
Acc.+Gyro.	91.61%(±2.60%)	92.50%(±4.03%)	90.87%(±4.06%)	90.64%(±4.16%)

Note: Acc. is Accelerometer data and Gryo. is Gyroscope data. (↑) and (↓) indicate higher and lower values are better for these metrics, respectively.

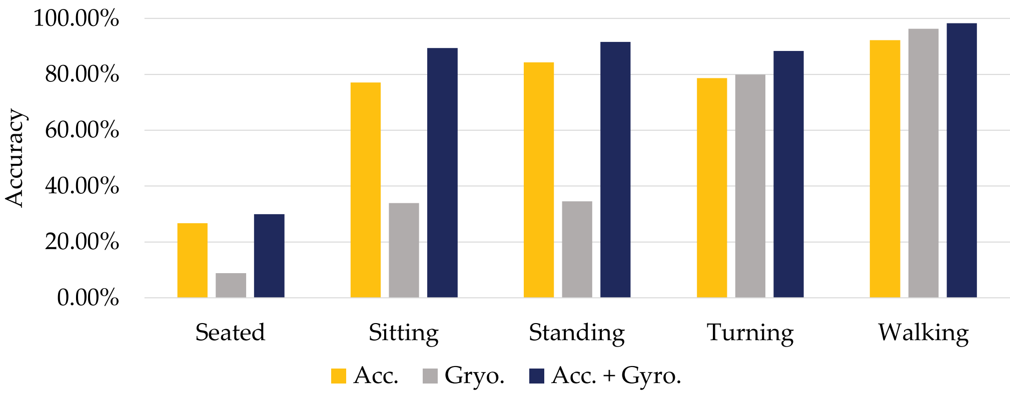


Fig. 5. User identification performance comparison of models trained and tested on different smartwatch sensor data types.

The combination of accelerometer and gyroscope data in Tables 6 and 7 showcases superior performance, reinforcing the effectiveness of amalgamating data from multiple sensors for user recognition.

5. Discussion

5.1. Effects of different types of sensors

To understand how various sensor types affect the performance of identification models, we examine the results presented in Tables 3 to 7. Fig. 5 depicts the models' performance when trained and assessed using different sensor types (accelerometer, gyroscope, and combination) from a smartwatch for user identification. The results are presented as average accuracy scores for five activities: seated, standing, turning, and walking.

These results highlight the substantial influence of sensor selection on user identification accuracy. Combining data from multiple sensors, like accelerometer and gyroscope, yields better performance than using a single sensor type. The effectiveness of the model varies depending on the activity type. Activities involving movement – turning and walking – achieve higher accuracy than static activities like seated, sitting, and standing. Relying solely on gyroscope data may not suffice for accurate user identification across all activities, mainly stationary ones. However, when combined with accelerometer data, it enhances overall performance.

5.2. Effects of different activity types

Based on the findings presented in Tables 3 to 7, it can be deduced that the performance of user identification models utilizing wearable device sensor data is significantly influenced by the type of activity undertaken. The primary observations suggest that dynamic activities, such as walking and rotational movements, achieve higher accuracy and F1-scores than static activities like sitting, remaining seated, or standing still.

The proposed Att-ResBiLSTM model consistently outperforms previous deep learning models across all activity categories when accelerometer and gyroscope data are amalgamated. Leveraging accelerometer and gyroscope data yields superior performance compared to relying solely on a single sensor data type, irrespective of the activity's nature. Identifying users during sedentary activities proves particularly challenging, as all models consistently underperform compared to other activity types.

These results indicate that when developing user identification models based on wearable device sensor data, it is crucial to consider the specific activity being performed and employ the most appropriate sensor modalities. Engaging in dynamic activities and utilizing both accelerometer and gyroscope data is more likely to enhance accuracy and F1-scores for user identification.

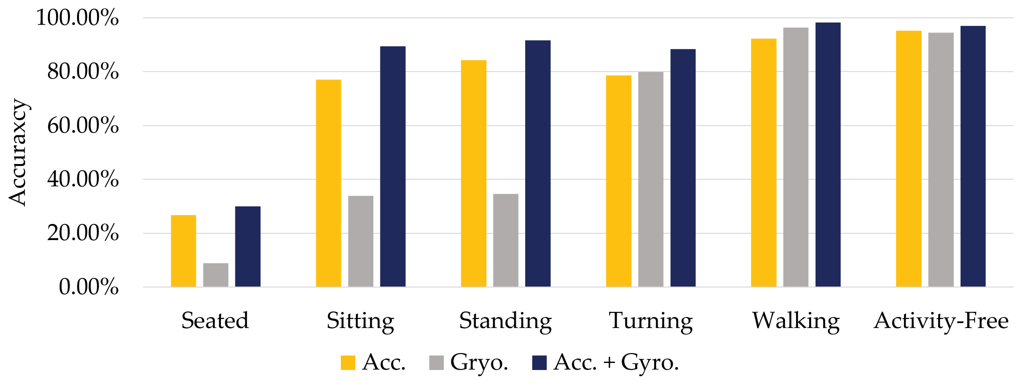


Fig. 6. Comparison results of the Att-ResBiLSTM using different types of activities and activity-free identification.

5.3. Activity-free user identification

Activity-free user identification is a biometric method not requiring individuals to engage in specific activities to validate their identity [39]. This approach does not necessitate or assume that users must perform specific movements or actions for the system to recognize them. For instance, people are not obligated to execute particular gestures or actions consciously. Consequently, individuals can carry out their daily routines and interact with mobile devices without altering or restricting their behavior.

This research evaluates the efficacy of the Att-ResBiLSTM system in identifying users based on sensor data from a comprehensive dataset without requiring specific behaviors. Our comparison, as illustrated by the findings in Fig. 6, demonstrates the efficiency of the proposed approach for identifying users across various activity types without the need for predefined activities.

Fig. 6 presents a comparative study highlighting the efficacy of the proposed Att-ResBiLSTM model in identifying users during activity-free scenarios. The model's performance in such instances is comparable to its performance in dynamic activities like walking and turning movements. This finding underscores the Att-ResBiLSTM model's capability to accurately identify users without relying on specific activity patterns. The discovery emphasizes the model's ability to discretely and precisely recognize users in real-world situations, even when they are not engaged in predefined activities.

5.4. Comparison with other deep learning models

In this section, we performed further investigations to assess the identification accuracy of the presented Att-ResBiLSTM model compared to other well-established models using deep learning. We have chosen five models that are extensively employed in diverse sequence modeling applications, such as HAR [40,41] and user identification [1].

- CNNs are highly efficient in extracting spatial characteristics from sensor input and have demonstrated impressive performance in human behavior identification challenges.
- LSTM is a specific type of RNN designed to identify long-term dependencies in time series data. LSTMs are particularly effective for modeling temporal patterns in sensor data.
- A BiLSTM is an extension of LSTMs that processes the input sequence in both forward and backward directions. This allows BiLSTMs to gather more context and enhance the effectiveness of sequence modeling.
- The gated recurrent unit (GRU) network is a type of RNN with a less complex structure than LSTM networks. However, GRUs can still accurately represent the relationships between events in a sequence across time.
- Bidirectional GRU (BiGRU): Similar to BiLSTM, BiGRU models handle the input sequence in both forward and backward directions to collect a more complete range of temporal data.

These models were selected based on their varied range of topologies and sequencing modeling potential, allowing for a comprehensive comparison with our proposed Att-ResBiLSTM model.

We established the assessment using the smartwatch sensor data of walking actions from the SP-SW dataset. The selection of walking activity data is based on its demonstrated ability to exhibit distinct and intricate patterns for each individual [29,42], rendering it an appropriate option for assessing user identification accuracy. Walking actions encompass dynamic movements that can record unique features of an individual's stride, arm swing, and other movement patterns, which can be utilized for identification reasons.

The comparison findings shown in Table 8 indicate that the recommended Att-ResBiLSTM model surpasses other advanced deep learning models in user identification through smartwatch sensor data from walking operations. It achieves the highest accuracy (98.29%) and F1-score (98.24%) when integrating accelerometer and gyroscope data. The exceptional effectiveness of Att-ResBiLSTM could be explained by its design, which adeptly collects spatial characteristics, temporal interdependencies, and attentive data, emphasizing its prospects for security-sensitive solutions like user authentication based on smartwatches.

Table 8
Comparative analysis with other deep learning models.

Model	Identification performance (%mean(±Std.))					
	Acc.		Gyro.		Acc.+Gyro.	
	Accuracy	F1-score	Accuracy	F1-score	Accuracy	F1-score
CNN	48.04%	45.89%	62.21%	61.73%	71.25%	71.26%
LSTM	61.71%	60.72%	48.96%	46.28%	77.25%	76.87%
BiLSTM	67.62%	66.49%	59.96%	58.60%	81.86%	81.90%
GRU	64.67%	63.56%	47.75%	45.59%	79.46%	78.89%
BiGRU	70.17%	69.08%	54.92%	53.74%	82.21%	81.96%
Att-ResBiLSTM	92.25%	92.24%	96.33%	96.36%	98.29%	98.24%

Note: Acc. is Accelerometer data and Gryo. is Gyroscope data.

Table 9
Mean ranks of models across different sensor datasets based on classification accuracy.

Sensor	Mean rank of each model for each sub-dataset		
	Acc.	Gyro.	Acc. + Gyro.
CNN	6.0	2.3	5.9
LSTM	4.1	5.4	4.6
BiLSTM	4.4	2.8	2.6
GRU	3.2	5.6	4.3
BiGRU	2.3	3.9	2.6
Att-ResBiLSTM	1.0	1.0	1.0

Note: Acc. is Accelerometer data and Gryo. is Gyroscope data.

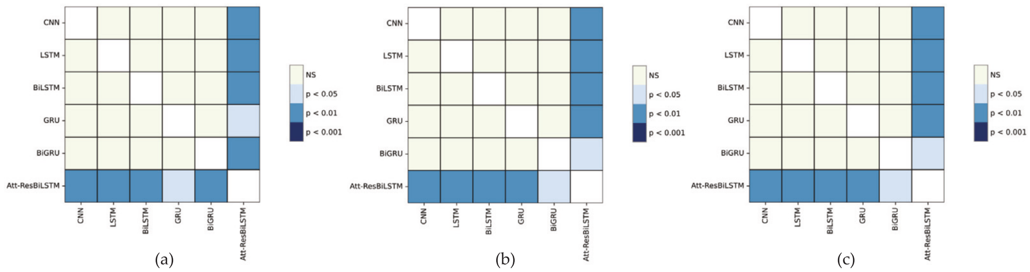


Fig. 7. Pairwise Nemenyi score plot generated for different sensor dataset: (a) accelerometer data, (b) gyroscope data, and (c) accelerometer and gyroscope data.

5.5. Statistical tests

We performed non-parametric statistical analyses to demonstrate the superior standing and efficacy of our suggested Att-ResBiLSTM model compared to the standard models (CNN, LSTM, BiLSTM, GRU, and BiGRU). We selected non-parametric tests because of their resilience when the underlying data distribution is unknown or non-normal. Initially, we conducted the Friedman test [43], which facilitates the comparison of numerous models across various datasets. Our null hypothesis posited that all models exhibit comparable effectiveness, with any observed variances attributable solely to random fluctuation. We randomly selected ten distinct subsets for the Friedman test, each comprising 230 samples from the test data for every evaluated sub-dataset, assuring uniform representation of all class labels. Subsequently, we assessed the classification accuracies of each model across these samples and ordered them correspondingly. The average rank for each model was determined following the formula:

$$R_j = \frac{1}{N}\sum_{i=1}^{N} r_j^i \tag{5}$$

Hence, r_j^i denotes the rank of the jth classifier or model for the ith sample. The computed mean ranks of the classifiers are presented in Table 9.

The Friedman test indicated statistically substantial variations in the effectiveness of the models. Since the p-value obtained was below the standard significance threshold of 0.05, we rejected the null hypothesis, asserting comparable outcomes among all models. This finding indicates that at least one model's effectiveness significantly differs. We performed a post-hoc analysis utilizing the Nemenyi test [44] to determine which models exhibit significant differences. Fig. 7 presents the pairwise Nemenyi scores (p-values) via heatmaps for each sensor dataset, visually depicting the effectiveness variations among models.

In these heatmaps, darker colors indicate lower p-values, suggesting more significant differences between the corresponding models. As observed in Fig. 7, the pairwise p-values between our proposed Att-ResBiLSTM model and the other baseline models

Table 10
Impact of the convolution block.

Model	Identification performance (%mean(±Std.))			
	Accuracy	Precision	Recall	F1-score
The Model without Convolution Block	96.82%(±0.83%)	96.97%(±0.79%)	96.80%(±0.85%)	96.83%(±0.83%)
Att-ResBiLSTM	96.98%(±0.89%)	97.11%(±0.85%)	96.93%(±0.91%)	96.97%(±0.90%)

Table 11
Effect the ResBiLSTM block.

Model	Identification performance (%mean(±Std.))			
	Accuracy	Precision	Recall	F1-score
Model without residual BiLSTM block	96.65%(±0.42%)	96.76%(±0.34%)	96.68%(±0.51%)	96.67%(±0.44%)
Att-ResBiLSTM	96.98%(±0.89%)	97.11%(±0.85%)	96.93%(±0.91%)	96.97%(±0.90%)

are consistently below the significance level of 0.05 across all sensor datasets. This result strongly supports the conclusion that our proposed Att-ResBiLSTM model significantly outperforms the five baseline models, thereby verifying its effectiveness and superiority in this classification task.

5.6. Ablation studies

Ablation studies are frequently employed in the domain of neural networks to assess the effectiveness of a model by investigating the impact of modifying particular elements [45,46]. Hence, we examine the consequences of ablation on the model we suggest through three research instances, whereby we modify several blocks and layers to assess the influence of these components on the proposed design [47]. This section analyzes the influence of various blocks and layers in our suggested Att-ResBiLSTM structure on the model's effectiveness. By implementing these ablation investigations, we aim to determine the most effective setup of our suggested model that produces the best level of identifying effectiveness.

5.6.1. Impact of the convolution block

In order to examine the influence of the convolutional block on the model's effectiveness, we carried out ablation research utilizing the SP-SW dataset. A standard model was generated by excluding the convolutional block from the Att-ResBiLSTM design but retaining the ResBiLSTM and attention layer. The standard model inputs the raw sensor data straight into the ResBiLSTM block without any spatial feature extraction.

Table 10 presents the significant findings of our ablation investigation. The original model, without the convolutional block, demonstrates inferior recognition effectiveness compared to the complete Att-ResBiLSTM model. The removal of the convolutional block leads to a decrease in the F1-score from 96.97% to 96.83%, underscoring the importance of spatial feature extraction for precise user identification. This discovery underscores the vital role of the convolutional block in capturing significant spatial features from the sensor input, thereby enhancing the overall effectiveness of the Att-ResBiLSTM model.

5.6.2. Impact of the ResBiLSTM block

To evaluate the effectiveness of the convolutional block in capturing spatial features in time series data, we performed an additional ablation experiment on our proposed Att-ResBiLSTM model. This experiment used a modified version of the Att-ResBiLSTM architecture as the baseline, which excluded the residual BiLSTM block. The results in Table 11 indicate that the proposed Att-ResBiLSTM model, incorporating the ResBiLSTM block, outperformed the baseline in identification tasks, increasing the F1-score from 96.67% to 96.97%.

5.6.3. Effect of the attention block

The Att-ResBiLSTM incorporates an attention mechanism that allows for selective concentration on significant areas within the acquired spatial–temporal feature representations prior to the ultimate classification process. Our hypothesis suggests that selectively focusing on essential signals and reducing the impact of irrelevant noise is crucial for achieving position-invariance. In order to assess the impact of the attention block, a modified version was evaluated where the attention block was removed while keeping the convolutional and ResBiLSTM modules intact.

As shown in Table 12, the ablation outcomes indicate that the suggested Att-ResBiLSTM model with the attention block exhibits superior identification effectiveness compared to the model without the attention block. When the attention block is incorporated into the overall structure, the F1-score rises from 96.44% to 96.97%, confirming its significance. This discovery implies that the attention mechanism significantly improves the model's capacity to concentrate on the most valuable characteristics and reduce interference, enhancing user identification accuracy.

The Att-ResBiLSTM model efficiently captures the most significant characteristics from wearable sensor data by utilizing convolutional blocks, a ResBiLSTM block, and an attention block.

Table 12
Effect the attention block.

Model	Identification performance (%mean(±Std.))			
	Accuracy	Precision	Recall	F1-score
The model without attention layer	96.40%(±0.42%)	96.64%(±0.35%)	96.41%(±0.40%)	96.44%(±0.38%)
Att-ResBiLSTM	96.98%(±0.89%)	97.11%(±0.85%)	96.93%(±0.91%)	96.97%(±0.90%)

5.7. Limitations

Although our suggested Att-ResBiLSTM model shows promising outcomes for user identification using smartwatch sensor data, it is crucial to recognize and address our method's constraints. The limitations encompass restrictions in the dataset, constraints in the sensors, and the model's ability to carry out individuals and actions that have yet to be previously encountered.

5.7.1. Dataset constraints

The SP-SW HAR dataset utilized in this investigation, despite its widespread usage and free accessibility, possesses specific limitations. The dataset comprises sensor data obtained from a restricted sample size of subjects (23 participants) falling within a defined population age range (23 to 66 years old). Although this dataset is a useful initial reference for assessing our suggested approach, it is necessary to conduct additional validation on more extensive and varied datasets to confirm the model's effectiveness and applicability. To enhance the model's robustness and applicability to a broader range of individuals, it would be beneficial to expand the number of people involved and include a more diverse variety of age groups, genders, and demographic characteristics.

5.7.2. Sensor limitations

The data used in our research was obtained from the accelerometer and gyroscope sensors incorporated into smartwatches. Although these sensors offer significant data for user identification, they possess specific constraints. The accuracy of accelerometer and gyroscope measurements can be influenced by variables such as the sensors' positioning, the instrument's orientation, and any disturbances caused by movement. Furthermore, the caliber and precision of sensor data can fluctuate among various smartwatch models and manufacturers. Further investigation is needed to examine how sensor positioning, alignment, and diversity affect the model's effectiveness. Furthermore, including data from supplementary sensors, such as monitors for heart rate or body temperature sensors, can substantially improve user identification accuracy.

5.7.3. Unseen users and activities

It is crucial to consider the limitations of the Att-ResBiLSTM model's effectiveness when applied to individuals and activities that have not been previously encountered. Although our model has impressive accuracy and an F1-score for particular operations and individuals in the SP-SW HAR dataset, its effectiveness with individuals and activities that have not been previously encountered still needs to be assessed. The model's capacity to extrapolate to unfamiliar users and discern patterns from previous unobserved actions is essential for practical applications. Subsequent investigations should prioritize acquiring data from a more extensive cohort of individuals and including a broader spectrum of actions to evaluate the model's efficacy in a more comprehensive array of genuine situations. Furthermore, the model's flexibility to unseen individuals and activities could be enhanced by investigating transfer learning and few-shot learning approaches.

6. Conclusion and future works

Conclusively, this research introduces an innovative hybrid residual deep learning network named Att-ResBiLSTM. It is designed to identify users on smartwatches by analyzing activity patterns. The proposed model combines convolutional layers, a ResBiLSTM component, and an attention mechanism to extract spatial features, capture long-term relationships in time series data, and improve the final recognition features. The user identification framework, based on smartwatches and using the Att-ResBiLSTM model, automates all steps of user identification by analyzing human activities.

The Att-ResBiLSTM model underwent evaluation utilizing the publicly accessible SP-SW HAR dataset, which includes sensor data from diverse human activities. The experimental findings indicate that the suggested model surpasses previous baseline deep learning models, with an accuracy of 98.29% and the highest F1-score of 98.24%. Furthermore, a comparison investigation showed that integrating accelerometer and gyroscope data results in enhanced performance in identifying users beyond the use of separate sensor modalities. The suggested technique provides a safe and fast way to identify users using smartwatches, improving the overall security of wearable devices. This approach utilizes activity patterns and a hybrid residual deep learning network to authenticate individual identities with high reliability and precision.

Future research endeavors may concentrate on verifying the proposed Att-ResBiLSTM model using supplementary datasets that include a wide range of physical activity patterns. Enhanced efficiency may be attained by creating more intricate and lightweight deep learning structures and investigating novel data representations using time–frequency analysis. Furthermore, examining the model's efficacy in real-life situations and its capacity to adjust to variations in user behavior over time would be beneficial for practical application.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Sakorn Mekruksavanich reports financial support was provided by Thailand Science Research and Innovation Fund (Fundamental Fund 2025). Anuchit Jitpattanakul reports financial support was provided by National Science, Research and Innovation (NSR). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by University of Phayao, Thailand; Thailand Science Research and Innovation Fund (Fundamental Fund 2025); National Science, Research and Innovation (NSRF); and King Mongkut's University of Technology North Bangkok, Thailand under Contract No. KMUTNB-FF-68-B-02.

Data availability

The original data presented in the study are openly available for the sp-sw-har-dataset of the "Dataset of inertial measurements of smartphones and smartwatches for human activity recognition" at "https://github.com/GeoTecINIT/sp-sw-har-dataset".

References

[1] Al-Mahadeen E, Alghamdi M, Tarawneh AS, Alrowaily MA, Alrashidi M, Alkhazi IS, et al. Smartphone user identification/Authentication using accelerometer and gyroscope data. Sustainability 2023;15(13). http://dx.doi.org/10.3390/su151310456.
[2] Meng Q, Lu P, Zhu S. A smartphone-enabled IoT system for vibration and noise monitoring of rail transit. IEEE Internet Things J 2023;10(10):8907–17. http://dx.doi.org/10.1109/JIOT.2022.3233051.
[3] Mekruksavanich S, Jitpattanakul A. Deep residual network for smartwatch-based user identification through complex hand movements. Sensors 2022;22(8). http://dx.doi.org/10.3390/s22083094.
[4] Alrawili R, AlQahtani AAS, Khan MK. Comprehensive survey: Biometric user authentication application, evaluation, and discussion. Comput Electr Eng 2024;119:109485. http://dx.doi.org/10.1016/j.compeleceng.2024.109485.
[5] Peinado-Contreras A, Munoz-Organero M. Gait-based identification using deep recurrent neural networks and acceleration patterns. Sensors 2020;20(23). http://dx.doi.org/10.3390/s20236900.
[6] De Marsico M, Mecca A, Barra S. Walking in a smart city: Investigating the gait stabilization effect for biometric recognition via wearable sensors. Comput Electr Eng 2019;80:106501. http://dx.doi.org/10.1016/j.compeleceng.2019.106501.
[7] Bajaber A, Elrefaei L. Biometric template protection for dynamic touch gestures based on fuzzy commitment scheme and deep learning. Mathematics 2022;10(3). http://dx.doi.org/10.3390/math10030362.
[8] Agrawal M, Mehrotra P, Kumar R, Shah RR. GANTouch: An attack-resilient framework for touch-based continuous authentication system. IEEE Trans Biometrics Behav Identity Sci 2022;4(4):533–43. http://dx.doi.org/10.1109/TBIOM.2022.3206321.
[9] Gadaleta M, Rossi M. IDNet: Smartphone-based gait recognition with convolutional neural networks. Pattern Recognit 2018;74:25–37. http://dx.doi.org/10.1016/j.patcog.2017.09.005.
[10] Giorgi G, Saracino A, Martinelli F. Using recurrent neural networks for continuous authentication through gait analysis. Pattern Recognit Lett 2021;147:157–63. http://dx.doi.org/10.1016/j.patrec.2021.03.010.
[11] Saini BS, Kaur N, Bhatia KS. Authenticating mobile phone user using keystroke dynamics. Int J Comput Sci Eng 2018;6:372–7. http://dx.doi.org/10.26438/ijcse/v6i12.372377.
[12] De Luca A, Hang A, Brudy F, Lindner C, Hussmann H. Touch me once and I know it's you! implicit authentication based on touch screen patterns. In: Proceedings of the SIGCHI conference on human factors in computing systems. New York, NY, USA: Association for Computing Machinery; 2012, p. 987–96. http://dx.doi.org/10.1145/2207676.2208544.
[13] Sae-Bae N, Memon N, Isbister K. Investigating multi-touch gestures as a novel biometric modality. In: 2012 IEEE fifth international conference on biometrics: theory, applications and systems. 2012, p. 156–61. http://dx.doi.org/10.1109/BTAS.2012.6374571.
[14] Frank M, Biedert R, Ma E, Martinovic I, Song D. Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication. IEEE Trans Inf Forensics Secur 2013;8(1):136–48. http://dx.doi.org/10.1109/TIFS.2012.2225048.
[15] Rocha CC, Lima JCD, Dantas MAR, Augustin I. A2BeST: An adaptive authentication service based on mobile user's behavior and spatio-temporal context. In: 2011 IEEE symposium on computers and communications. 2011, p. 771–4. http://dx.doi.org/10.1109/ISCC.2011.5983933.
[16] Sabharwal M. Multi-modal biometric authentication and secure transaction operation framework for E-banking. Int J Bus Data Commun Netw 2017;13(1):102–16. http://dx.doi.org/10.4018/IJBDCN.2017010109.
[17] Jakobsson M, Shi E, Golle P, Chow R. Implicit authentication for mobile devices. In: Proceedings of the 4th USENIX conference on hot topics in security. USA: USENIX Association; 2009, p. 9.
[18] Casale P, Pujol O, Radeva P. Personalization and user verification in wearable systems using biometric walking patterns. Pers Ubiquitous Comput 2011;16:563–80.
[19] Hinton GE, Osindero S, Teh Y-W. A Fast Learning Algorithm for Deep Belief Nets. Neural Comput 2006;18(7):1527–54. http://dx.doi.org/10.1162/neco.2006.18.7.1527.
[20] Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L. Large-scale video classification with convolutional neural networks. In: 2014 IEEE conference on computer vision and pattern recognition. 2014, p. 1725–32. http://dx.doi.org/10.1109/CVPR.2014.223.
[21] Taigman Y, Yang M, Ranzato M, Wolf L. DeepFace: Closing the gap to human-level performance in face verification. In: 2014 IEEE conference on computer vision and pattern recognition. 2014, p. 1701–8. http://dx.doi.org/10.1109/CVPR.2014.220.
[22] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: 2018 IEEE/CVF conference on computer vision and pattern recognition. 2018, p. 7132–41. http://dx.doi.org/10.1109/CVPR.2018.00745.
[23] Chen L, Zhang H, Xiao J, Nie L, Shao J, Liu W, et al. SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning. In: 2017 IEEE conference on computer vision and pattern recognition. Los Alamitos, CA, USA: IEEE Computer Society; 2017, p. 6298–306. http://dx.doi.org/10.1109/CVPR.2017.667.

[24] Dai J, Qi H, Xiong Y, Li Y, Zhang G, Hu H, et al. Deformable convolutional networks. In: 2017 IEEE international conference on computer vision. 2017, p. 764–73. http://dx.doi.org/10.1109/ICCV.2017.89.
[25] Wu Y, He K. Group normalization. Int J Comput Vis 2018;128:42–755.
[26] Yang J, Ren Z, Gan C, Zhu H, Parikh D. Cross-channel communication networks. In: Advances in neural information processing systems, vol. 32, Curran Associates, Inc.; 2019.
[27] Benegui C, Ionescu RT. Convolutional neural networks for user identification based on motion sensors represented as images. IEEE Access 2020;8:61255–66. http://dx.doi.org/10.1109/ACCESS.2020.2984214.
[28] Angrisano A, Bernardi ML, Cimitile M, Gaglione S, Vultaggio M. Identification of walker identity using smartphone sensors: An experiment using ensemble learning. IEEE Access 2020;8:27435–47. http://dx.doi.org/10.1109/ACCESS.2020.2971693.
[29] Weiss GM, Yoneda K, Hayajneh T. Smartphone and smartwatch-based biometrics using activities of daily living. IEEE Access 2019;7:133190–202. http://dx.doi.org/10.1109/ACCESS.2019.2940729.
[30] Musale P, Baek D, Werellagama N, Woo SS, Choi BJ. You walk, we authenticate: Lightweight seamless authentication based on gait in wearable IoT systems. IEEE Access 2019;7:37883–95. http://dx.doi.org/10.1109/ACCESS.2019.2906663.
[31] Ahmad M, Alqarni M, Khan A, Khan A, Chauhdary S, Mazzara M, et al. Smartwatch-based legitimate user identification for cloud-based secure services. Mob Inf Syst 2018;2018:1–14. http://dx.doi.org/10.1155/2018/5107024.
[32] Neverova N, Wolf C, Lacey G, Fridman L, Chandra D, Barbello B, et al. Learning human identity from motion patterns. IEEE Access 2016;4:1810–20. http://dx.doi.org/10.1109/ACCESS.2016.2557846.
[33] Matey-Sanz M, Casteleyn S, Granell C. Dataset of inertial measurements of smartphones and smartwatches for human activity recognition. Data Brief 2023;51:109809. http://dx.doi.org/10.1016/j.dib.2023.109809.
[34] Mostayed A, Kim S, Mazumder MMG, Park SJ. Foot step based person identification using histogram similarity and wavelet decomposition. In: 2008 international conference on information security and assurance. 2008, p. 307–11. http://dx.doi.org/10.1109/ISA.2008.46.
[35] Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL. A public domain dataset for human activity recognition using smartphones. In: The European symposium on artificial neural networks. 2013, p. 437–42.
[36] Kwapisz JR, Weiss GM, Moore SA. Activity recognition using cell phone accelerometers. SIGKDD Explor Newsl 2011;12(2):74–82. http://dx.doi.org/10.1145/1964897.1964918.
[37] Kingma DP, Ba J. Adam: A method for stochastic optimization. 2017, arXiv:1412.6980.
[38] Janocha K, Czarnecki WM. On loss functions for deep neural networks in classification. 2017, arXiv:1702.05659.
[39] Sanchez Guinea A, Heinrich S, Mühlhäuser M. Activity-free user identification using wearables based on vision techniques. Sensors 2022;22(19). http://dx.doi.org/10.3390/s22197368.
[40] Abbaspour S, Fotouhi F, Sedaghatbaf A, Fotouhi H, Vahabi M, Linden M. A comparative analysis of hybrid deep learning models for human activity recognition. Sensors 2020;20(19). http://dx.doi.org/10.3390/s20195707.
[41] Kaseris M, Kostavelis I, Malassiotis S. A comprehensive survey on deep learning methods in human activity recognition. Mach Learn Knowl Extr 2024;6(2):842–76. http://dx.doi.org/10.3390/make6020040.
[42] Mekruksavanich S, Jitpattanakul A. Biometric user identification based on human activity recognition using wearable sensors: An experiment using deep learning models. Electronics 2021;10(3). http://dx.doi.org/10.3390/electronics10030308.
[43] Singh PK, Sarkar R, Nasipuri M. Significance of non-parametric statistical tests for comparison of classifiers over multiple datasets. Int J Comput Sci Math 2016;7(5):410–42. http://dx.doi.org/10.1504/IJCSM.2016.080073.
[44] Singh P, Sarkar R, Nasipuri M. Statistical validation of multiple classifiers over multiple datasets in the field of pattern recognition. Int J Appl Pattern Recognit 2015;2(1):1–23. http://dx.doi.org/10.1504/IJAPR.2015.068929.
[45] Montaha S, Azam S, Rafid AKMRH, Ghosh P, Hasan MZ, Jonkman M, et al. BreastNet18: A high accuracy fine-tuned VGG16 model evaluated using ablation study for diagnosing breast cancer from enhanced mammography images. Biology 2021;10(12). http://dx.doi.org/10.3390/biology10121347.
[46] de Vente C, Boulogne LH, Venkadesh KV, Sital C, Lessmann N, Jacobs C, et al. Improving automated COVID-19 grading with convolutional neural networks in computed tomography scans: An ablation study. 2020, arXiv:2009.09725.
[47] Meyes R, Lu M, de Puiseau CW, Meisen T. Ablation studies in artificial neural networks. 2019, arXiv:1901.08644.