IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 11, 2021, accepted June 21, 2021, date of publication June 25, 2021, date of current version July 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3092435

A Deep Learning-Based Human Identification System With Wi-Fi CSI Data Augmentation

HYUNGGEUN MO AND SEUNGKU KIM
Department of Electronics Engineering, Chungbuk National University, Cheongju-si 28644, South Korea

Corresponding author: Seungku Kim (kimsk@cbnu.ac.kr)

This work was supported by the National Research Foundation of Korea (NRF) through the Korea Government (MSIP) under Grant 2019R1F1A1061970.

This work did not involve human subjects or animals in its research.

ABSTRACT Human identification systems generally include face recognition, iris recognition, radio frequency identification tags, and fingerprint recognition systems. However, these systems pose problems such as privacy violations, loss concerns, lighting requirements, and additional installation costs. Several studies have been conducted on human identification systems using Wi-Fi signals to address these problems. However, there exist problems such as a low number of identified per-sons, low accuracy, and high cost of data collection. In this paper, we present a deep-learning-based human identification system via Wi-Fi channel state information. To reduce the cost of data collection and increase the accuracy of human identification, we propose a data preprocessing and data augmentation process. They achieve an accuracy improvement of approximately 7%. In addition, we implemented one machine learning model and three deep learning models and demonstrated that the CLSTM model is suitable for the application through performance evaluation. The proposed system can identify up to 8 subjects with an accuracy of about 92%.

INDEX TERMS Channel state information, data augmentation, deep learning, human identification, Wi-Fi.

I. INTRODUCTION

Internet of Things (IoT) is an intelligent technology that analyzes, predicts, and makes decisions regarding different objects based on sensor information, such as humidity, temperature, location, and communication status. Recently, artificial intelligence technologies have been widely applied, and IoT is rapidly developing in various fields. For instance, active research on machine learning and deep learning is underway in the fields of prediction and classification based on time series data collected by sensors as well as video and voice data. Hence, IoT has experienced an increase in demand and supply for drones, autonomous driving, medical services, smart cities, smart buildings, and smart farms.

As the spread and use of IoT devices increases, indoor access management and building security for smart buildings and smart homes have become essential services. For these services, a human identification system is indispensable. Many human identification systems have been studied in various ways. The representative commercial human identification systems include face recognition using a camera [1], Radio frequency identification (RFID) tags [2], iris recognition [3], and fingerprint recognition [4]. However, these systems exhibit some issues. Commonly, extra installation costs are required for these human identification systems. In addition, the systems of face recognition, iris recognition, and fingerprint recognition may infringe on privacy, and in case of RFID tags, the risk of loss exists.

In order to solve these problems, a human identification system using CSI of Wi-Fi has recently been studied. Wi-Fi is one of the IT technologies that is most prominent in our daily life and can be used in most indoor environments such as home, school, office, etc. Using this technology, we can easily and conveniently identify humans without extra installation costs, invasion of privacy, and risk of loss.

Initially, human identification via received signal strength indicator (RSSI) analysis of Wi-Fi signals has been proposed in [5], [6]. However, this is a very unstable indicator because it is heavily influenced by the communication environment such as the multi-path effect. Channel state

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

information (CSI) [7] provides physical layer information of the Wi-Fi devices. It is a more stable and detailed indicator than RSSI. By means of CSI, many challenging Wi-Fi signal-based applications such as indoor localization, gesture recognition, and human counting have been proposed [8].

Various Wi-Fi-based human identification approaches and systems have been developed. However, a major challenge has not been resolved yet. i.e., the perturbations of the Wi-Fi signals are subject to change with respect to the diversified motion speed and body types of individuals who conduct activities. In addition, a large amount of data is required to improve the accuracy of the model, and a large amount of time is required for data collection. In this paper, to collect a small amount of data and improve the generalization of the model according to the walking speed of a person, we propose a WIID applying data augmentation. WIID was designed with three deep learning models: convolutional neural network (CNN), long and short-term memory (LSTM), and convolutional long- and short-term memory (CLSTM) to evaluate human identification. Virtual CSI data is increased by applying data increasing techniques such as sliding window [9] and time warping [10]. The proposed data argument improved the deep learning accuracy by about 7%, and CLSTM showed an accuracy of 92% in a scenario with 8 subjects. This is by far the highest accuracy compared to those of the existing human identification techniques.

Our primary contributions are as follows:

- We proposed two data augmentation schemes, i.e., sliding window and time warping, to produce a meaningful dataset with a small amount of data. This improves the accuracy of the deep learning model and increases the practicality of the human identification system. The data augmentation technique has been first used in the Wi-Fi CSI-based human identification system.
- We designed deep learning models optimized for a small dataset. These solved the problems regarding overfitting and underfitting in the learning stage and derived the CLSTM that demonstrates the best performance for the human identification application.
- The proposed system was implemented on commodity Wi-Fi APs to collect Wi-Fi CSI dataset for a total of 8 subjects and to identify each subject. As a result of the experiment, the performance of the proposed system was improved where 7 or more subjects exist.

The remaining parts of this paper are structured as follows. Section 2 presents the trends and problem definition in related research. Section 3 introduces some technical backgrounds. Section 4 presents the proposed human identification system, and Section 5 evaluates the performance of the proposed system. The conclusions are then stated in Section 6.

II. RELATED WORKS

At the early stages of Wi-Fi signal-based research, RSSI-based statistical models have dominated. However, these statistical models achieved a low stability due to unpredictable multipath effects in indoor environments. It is known that CSI-based models are more effective and more reliable [15]. They can be used in indoor localization [11], [16], human counting system [12], gesture recognition [13], etc. Wi-Fi CSI-based applications using machine or deep learning models have shown higher accuracy than the statistical models.

Zhang *et al.* [15] initiated the human identification using Wi-Fi CSI. The author assumed that every person had a unique walking pattern. Every walking pattern generated a unique perturbation of each Wi-Fi CSI spectrum, which could be used to extract the unique feature of the person. It used a continuous wavelet transformation (CWT) preprocessing algorithm for extracting the signal of various frequency bands. The proposed identification scheme was implemented by sparse approximation-based classification (SAC) model and achieved an average accuracy of 77% and 93% in groups of two and six persons, respectively. The accuracy of this model depended on the number of subjects participating in the experiment. It showed 77% accuracy in an experiment with six subjects.

Zeng *et al.* [14] proposed a system named WiWho, which performed human identification by extracting step counts and gait patterns via Wi-Fi CSI. It used dynamic time warping (DTW) and Butterworth filter for preprocessing CSI data. The subject had to walk 1 meter parallel to the Wi-Fi APs. WiWho achieved an average accuracy of 92% and 80% in groups of three and six subjects, respectively. This resulted in only the results for a maximum of six subjects, and many data had to be collected for the experiment.

Nipu *et al.* [17] used decision trees and random forest models to identify persons. They choose a path with large CSI fluctuations in advance and collected data from that path. The average, maximum, minimum, skewness, and energy values of the collected data were applied to the model. The decision tree model achieved an average accuracy of 95% and 84% in groups of two and five subjects, respectively; whereas the random forest algorithm obtained an accuracy of 78% to 97.5%.

The conventional approaches for Wi-Fi CSI-based human identification have three common problems. First, all the proposed schemes present low accuracy. Second, very few subjects can be identified by their human identification schemes. All the studies so far have classified up to six subjects. Third, they require many CSI data collection for high accuracy to train their models. However, collecting CSI data should be minimized because CSI data collection requires a lot of cost and time. The proposed human identification system focuses on methods to solve these three problems.

III. TECHNICAL BACKGROUND

A. Wi-Fi CHANNEL STATE INFORMATION

Wi-Fi technology provides high throughput via the multiple-input multiple-output (MIMO) method and achieves high frequency efficiency via orthogonal frequency division multiplexing (OFDM). OFDM divides the entire spectrum of Wi-Fi into multiple orthogonal subcarriers, and each subcarrier is divided into narrowband channel sets through MIMO. Each subcarrier uses the same modulation and coding scheme for Wi-Fi data transmission. These features facilitate Wi-Fi everywhere.

Many researchers have presented new applications by means of Wi-Fi as sensors. In the early stage, research using Wi-Fi as sensors mainly used RSSI as data. Recent studies have focused on using CSI, which reflects the channel state in more detail, to collect ambient data. CSI is physical layer information that refers to know channel properties of each OFDM subcarrier. We can analyze the changes of propagated signal such as scattering, attenuation, diffraction, fading, and reflection between the transceiver and receiver.

In the case of a MIMO-OFDM channel with M transmit antennas, N receive antennas, and k subcarriers, the CSI matrix may be expressed as a 3D matrix representing the amplitude attenuation and phase shift of the multipath channel. That is, the matrix is composed of CSI elements of $M \times N \times k$ subcarriers. Hence, the CSI time series matrix is expressed in terms of time, frequency, and transmission/reception antenna pairs.

CSI data are mainly collected in two ways: the use of the 1) Intel 5300 network interface card (NIC) [18] and 2) Qualcomm Atheros series NIC. The number of subcarriers depends on the types of NIC and selected bandwidth. For example, in the 20 MHz bandwidth, the number of CSI subcarriers is 30 and 56 for Intel 5300 NIC and Qualcomm Atheros series NIC, respectively.

B. MULTIPATH EFFECT

Wi-Fi standards such as IEEE 802.11a/g/n/ac adopt OFDM for high bandwidth efficiency. OFDM divides a transmission frequency channel into several subcarriers with a fixed length. Each subcarrier uses the same modulation scheme at a low symbol rate. Every subcarrier of the channel is orthogonal to each other. This orthogonality allows to avoid inter-carrier interference.

Signals can be propagated to line of sight (LOS) and non-line of sight (NLOS). In NLOS, the propagated signals can be reflected or refracted to multiple paths by objects (e.g., walls, floors, ceilings, humans). OFDM is robust against frequency selective fading but sensitive to time selective fading in this multipath environment [19]. The CSI of each subcarrier knows the multipath effect through the temporal changes of amplitude and phase. By means of this characteristics, conventional works use Wi-Fi CSI as ambient sensor data.

We conducted an experiment to examine the difference in the multipath effect according to body shape, gait, and stride length of persons via CSI amplitude changes. Every subject walks near installed Wi-Fi devices. Fig. 1 shows CSI amplitude changes in a subcarrier by each subject. It was difficult to confirm the pattern of CSI amplitude changes, but the difference between the subjects was clear. We trained different CSI results for each subject in deep learning system and performed human identification.

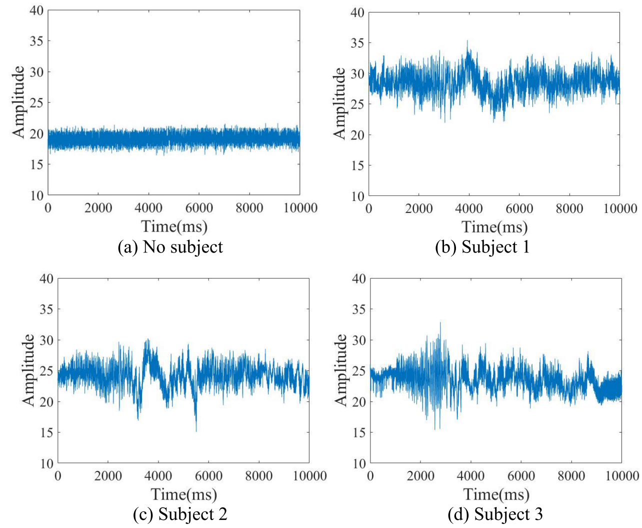


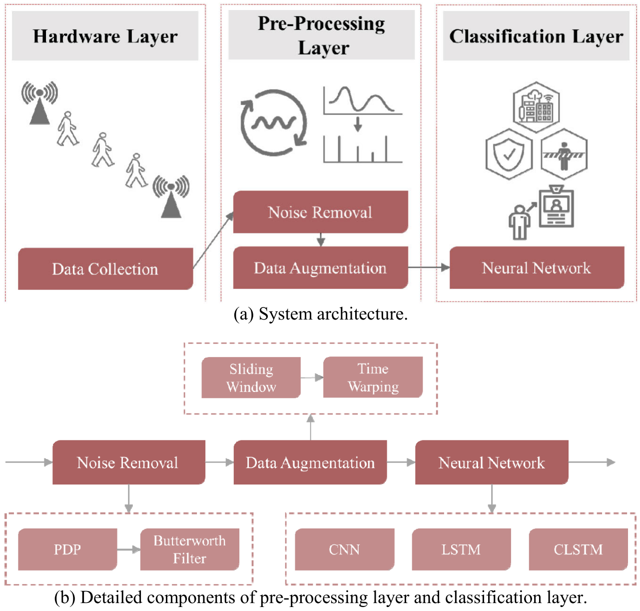
FIGURE 1. The changes in CSI amplitude for different subjects.

IV. SYSTEM DESIGN AND IMPLEMENTATION

Fig. 2 illustrates a deep learning-based human identification system. The system consisted of three layers: 1. hardware, 2. pre-processing, and 3. classification. The role of the hardware layer was to collect CSI data to investigate the multipath effect. The sampling rate for the data collection was set to 1,000 Hz. The processing layer removed the collected CSI data noise using channel impulse response (CIR) and Butterworth filter and performed data augmentation using sliding window and time warping schemes. The pre-processed data was inserted to the neural network for human identification in the classification layer. We implemented convolutional neural network (CNN), long short-term memory (LSTM), and convolutional long short-term memory (CLSTM) to classify the subjects.

A. HARDWARE LAYER

The proposed system operated in an indoor environment with a pair of commercial Wi-Fi modules. The receiver had NIC and was capable of collecting CSI data. In this paper, we used an Intel 5300 NIC supporting an IEEE 802.11n standard. The transmitter and receiver each used two MIMO antenna. The frequency band and bandwidth were 5.32 GHz and 20 MHz, respectively. The 20 MHz bandwidth was divided into 30 subcarriers; hence, each packet provided a $2 \times 2 \times 30$ CSI matrix. A transmitter sent one packet every millisecond and a receiver collected this in memory. Fig. 3 shows the experimental environment in which we collected the CSI data. In an 8×10 m2 classroom, we collected CSI data where the distance between the transmitter and receiver was 5.2 m. All subjects walked a dedicated path of 7.2 m at a speed of about 5 km/h. They waited at the starting point and walked to the end point when the start alarm went off. All subjects



(a) System architecture.

(b) Detailed components of pre-processing layer and classification layer.

FIGURE 2. **A deep learning-based human identification system.**

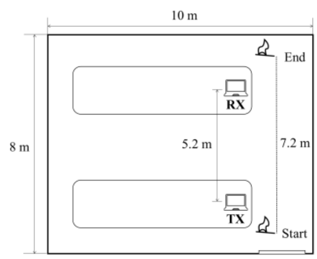


FIGURE 3. **Experimental environment.**

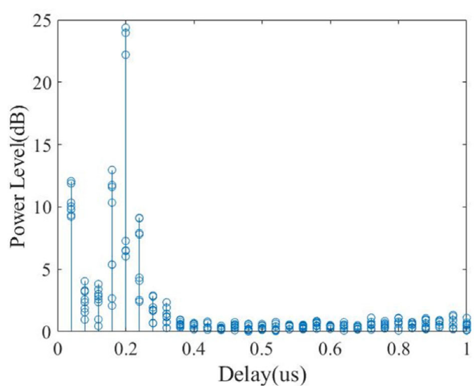
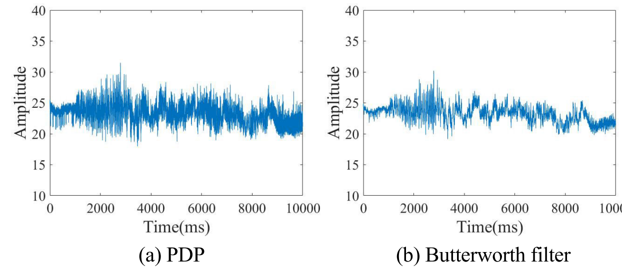


FIGURE 4. **Power delay profile of raw CSI measurements.**

arrived at the end point within 10 seconds. One dataset to train deep learning consisted of CSI matrices of packets for 10 seconds (e.g., 10,000 packets).

B. PROCESSING LAYER

Because OFDM is sensitive to time selective fading, the multipath effect caused delayed signals. Fig. 4 illustrates the power delay profile (PDP) [20] of one packet transmission. The PDP was derived by channel frequency response (CFR) via inverse fast Fourier transformation and norm. In the PDP,



(a) PDP (b) Butterworth filter

FIGURE 5. **Subject 3's results from the noise removal step.**

a packet contained a reflected or refracted delayed signal. Since we were only interested in the multipath signals by the subject, the signals delayed by the distance objects such as walls, ceilings, and obstacles were unnecessary. We only needed to include the multipath effect by the subject moving around the Wi-Fi devices. In [20], they proved that the maximum delayed multipath signal by the subject was within 0.5 μs. Therefore, we regarded the signals after 0.5 μs as the uncorrelated signals and removed them from the collected CSI data.

Fig. 5 (a) is the result of removing uncorrelated components from the CSI data in Fig. 1 (b). After removing the uncorrelated signals, the amplitude fluctuation was larger because only the multipath effect caused by the subject was reflected. This allowed the deep learning system to better analyze the subject's features.

Wi-Fi signal includes low frequency noise caused by diffraction, refraction, scattering, and penetration, and high frequency noise due to hardware. This causes the loss of accuracy to deep learning system. We used the Butterworth filter to remove these noises. Wang *et al.* [13] showed that the frequency ranges of all human activities caused CSI changes within 300 Hz and walking led to CSI changes of about 10 Hz. In this study, we removed noise in the frequency bands other than 5 to 100 Hz using the Butterworth filter. Fig. 5 (b) is the noise removal result from Fig. 5 (a) via the Butterworth filter. The features of CSI data were maintained even after passing through the Butterworth filter.

Deep learning models such as CNN and LSTM have been successfully applied to various classification tasks. To obtain high accuracy, they require large amounts of data. Several data augmentation algorithms such as jitter, scaling, cropping, and warping [9] have been used to enhance the performance of deep learning models. They generated probable data and maintained their labels. Many deep learning-based works [22]–[24] have used the data augmentation to improve accuracy, generalization, and prevent overfitting. [22] augments CSI spectrum data through nine different data transformation processes. This work aims at applying human behavior recognition, and CSI amplitude change is less than that of human identification. Moreover, the CSI change rate is not important because it does not consider the speed of human movement, i.e., temporal information. These data augmentation techniques are not suitable for the human identification application.

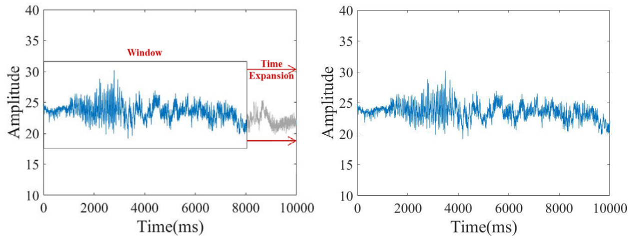


FIGURE 6. Subject 3's results from the data augmentation.

The human identification system that uses Wi-Fi has difficulty in collecting large datasets. The proposed human identification system uses two data augmentation algorithms such as sliding window and time warping to increase the collected CSI data. The sliding window algorithm inspired by computer vision and Transmission Control Protocol (TCP) is commonly used to augment time series data. The CSI matrix is a time series data, so the sliding window algorithm is a proper data augmentation for that. Human walking rhythm has repeated patterns, and CSI data holds the information of these patterns. To generate virtual datasets by using this information, the following sliding window algorithm was applied:

$$SW(\omega, \varepsilon) = \{W_0, W_\varepsilon, W_{2\varepsilon}, \cdots, W_{n-\omega}\} \tag{1}$$

Equation (1) represents time series data of each window for the sliding window algorithm where ω and ε are a defined window size and moving distance of the window, respectively. The collected CSI data is expressed as $T = \{t_1, t_2, t_3, \cdots, t_n\}$where n is the size of times series T, and $W_i = \{t_{i+1}, t_{i+2}, t_{i+3}, \cdots, t_{i-\omega}\}$ is the partial time series data divided by a fixed window size, ω. One original data set can be increased $(n - \omega/\varepsilon) + 1$ times through the sliding window algorithm. For example, 10,000 data (n) are generated by an original data set collected for 10 seconds. If the window size (ω) is selected as 8,000 and the moving distance (ε) is set to 200, additional data of 110,000 are produced.

The length of the virtual data generated by the sliding window was shorter than the length of the original data. We adjusted the virtual data to the length of the original data for the deep learning train. A time warping algorithm could expand or compress the time series data. By means of this algorithm, the size of the virtual data was expanded to the size of the original data shown in Fig. 6. The time warping also served to diversify a subjects' walking speed through data expansion and compression. We added data to increase and decrease the walking speed through time warping.

By means of data augmentation, we obtained 19 virtual CSI datasets from one collected CSI dataset. That is, the size of the datasets for deep learning training grew 19 times. Through the sliding window with SW(8000, 200) and SW(9000, 200), we secured 11- and 6-times virtual data, respectively. In addition, original data for 11 and 12 seconds were compressed to 10 seconds via time warping to produce double virtual data. The datasets to be trained in each deep learning model consisted of 70 original datasets and 1,330 virtual datasets per subject. One dataset included 1,000 CSI data, so we used a total of 1,337,000 data for the deep learning train.

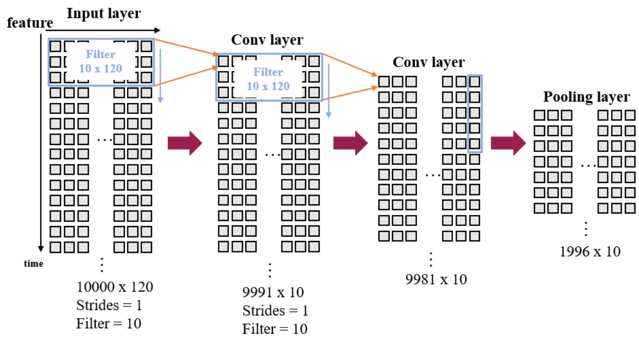


FIGURE 7. A proposed 1D-CNN architecture.

C. NEURAL NETWORK

In this study, we used three deep learning models: one-dimensional convolutional neural networks (1D-CNN), long short term (LSTM), and convolutional long-term memory (CLSTM). CNN is a deep neural network imitating the human optic nerves and is specialized in processing a grid data format. The Wi-Fi CSI matrix is 1D time series data, so we used 1D-CNN [25]. Fig. 7 shows a proposed 1D-CNN architecture. We use the convolution layers to extract hierarchical features from low level to high level. CSI data including a size of 10000×120 was input in the input layer, and convolution was performed through a 10×120 filter in the convolution layer. This enabled us to extract CSI features in increments of 0.001 seconds. Finally, a pooling layer was used to process the exponentially increased features by the CNN filter. In this study, we used 10 filters per convolution layer and set the stride to 1.

A CNN does not consider the temporal order of the data. However, since Wi-Fi CSI matrix is time series data, we investigated LSTM that was a special kind of recurrent neural network (RNN) to maintain temporal information in the CSI matrix. It is widely used for classification, processing, and prediction using time series data. The LSTM played an important role in processing time series data of human identification.

CLSTM is an ensemble model of CNN and LSTM. We stacked the outputs of the CNN layer across time as the input of the following LSTM layers. Fig. 8 shows a proposed human identification architecture including CLSTM network. The pre-processed CSI data is iput into the convolutional layer. The inputted data is feature extracted every 0.001 second by the convolution filter, and then sent to the next convolution layer. In the 1D-CNN layers, filters of 100×120 and 100×10 were used, respectively. Afterward, a new layer was obtained by Max pooling, and a time distributed layer was used to reduce the data dimension before LSTM input. Temporal features were input to the DNN layer

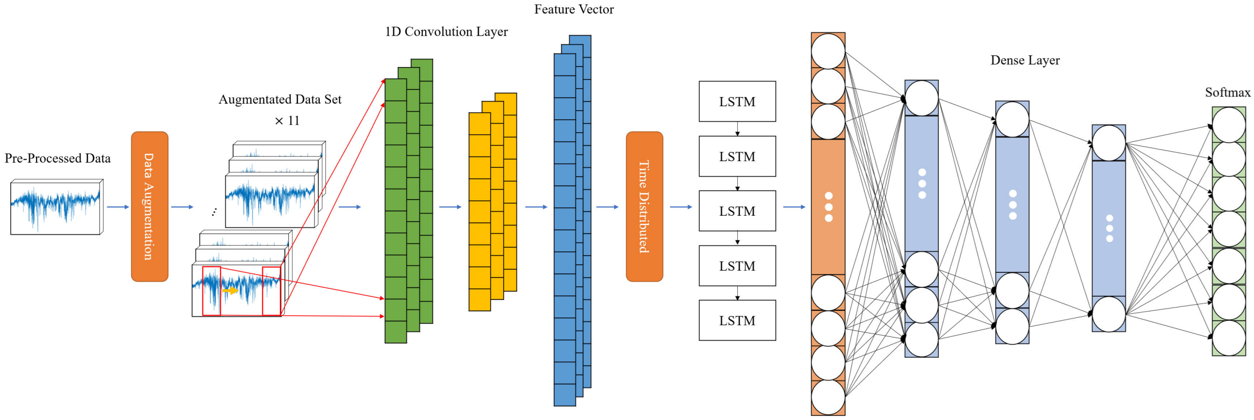


FIGURE 8. A proposed human identification architecture with CLSTM.

TABLE 1. Hyper parameters of the deep learning models.

Parameter	CNN	LSTM	CLSTM
Batch Size	128	32	128
Kernel Size	10	-	10
Strides	1	-	1
Activation	ReLu	ReLu	ReLu
Optimizer	Adam	Adam	Adam
Epochs	50	100	50
Loss Function	Categorical Cross-Entropy	Categorical Cross-Entropy	Categorical Cross-Entropy

through the LSTM layer, and data was finally classified through softmax.

V. PERFORMANCE EVALUATION

A. EXPERIMENTAL SETUP

We conducted the experiment that consist of none and eight subjects. One experiment was called a class, and up to nine classes were evaluated. The subjects consisted of five males and three females aged 22 to 28 years. Heights and weights were different for each subject. CSI matrices were collected through the hardware layer. The data of each subject was collected for 12 seconds while the subjects walked from the starting point to the end point (see Fig. 3). The walking path of subjects was located 1 m parallel to the transmitter and receiver. There were no restrictions on subjects' behaviors (e.g., using a smartphone, running, walking, jumping, etc.) when they walked. We collected 100 datasets for each class, of which 70 were used for training datasets and 30 for test datasets. The training datasets were augmented 19 times and 20% of them were randomly used for the validation. Table 1 presents the hyper parameters of the deep learning models.

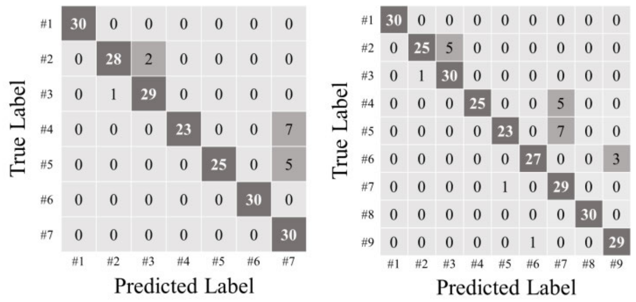
B. EXPERIMENTAL RESULTS

We focused on evaluating the performance regarding the effects of data augmentation and deep learning models. Most of the past studies have increased accuracy by increasing the amount of data collected or by reducing the number of classes to be classified. In the case of SVM, a machine learning algorithm that has been widely used in existing human identification, it was seen that the accuracy deteriorated rapidly as the number of classes increased. For nine classes, the accuracy was 66%. Papers using SVM derived results using only up to five classes. In the case of deep learning models, more people could be analyzed because the accuracy decreased according to the reduction of the number of classes.

TABLE 2. Accuracy by number of classes.

Models	Accuracy (%)			
	3-Class	5-Class	7-Class	9-Class
Without data augmentation				
SVM [26][27]	91	82	73	66
CNN	98	95	90	89
LSTM	94	89	81	77
CLSTM	98	96	83	85
With data augmentation				
CNN	98	96	92	91
LSTM	94	91	90	86
CLSTM (Proposed)	98	96	93	92

Table 2 shows the accuracy of the proposed models with and without the data augmentation technique. The accuracy of the SVM machine learning model decreased significantly as the number of classes to be classified increased. The nine-class identification clearly did not operate normally as it achieved an accuracy of 66%. For the deep learning model, the accuracy also decreased as the number of classes to be classified increased, but the decrease was smaller than that of the machine learning model. CNN without data augmentation achieved the highest accuracy in the four models. This result indicates that CNN works well as a feature extraction role from a lot of data. LSTM is an appropriate network for processing time series data, but it represents low performance compared with other two deep learning models because the size of the input data set is small, and the features of the input data are unclear. A small amount of data set does not reflect



(a) 7 classes (b) 9 classes

FIGURE 9. Confusion matrices of CLSTM.

TABLE 3. Evaluation parameters.

Models		Precision	Recall	F1-score	Accuracy
Without data augmentation					
3 Class	CNN	0.99	0.98	0.98	0.98
	LSTM	0.92	0.96	0.94	0.94
	CLSTM	0.99	0.96	0.97	0.98
5 Class	CNN	0.92	0.98	0.95	0.95
	LSTM	0.86	0.93	0.89	0.89
	CLSTM	0.93	0.97	0.95	0.96
7 Class	CNN	0.88	0.95	0.91	0.90
	LSTM	0.77	0.79	0.78	0.81
	CLSTM	0.85	0.79	0.82	0.83
9 Class	CNN	0.90	0.89	0.89	0.89
	LSTM	0.74	0.74	0.74	0.77
	CLSTM	0.71	0.92	0.89	0.85
With data augmentation					
3 Class	CNN	0.99	0.98	0.98	0.98
	LSTM	0.96	0.91	0.93	0.94
	CLSTM	0.99	0.98	0.98	0.98
5 Class	CNN	0.92	0.97	0.94	0.96
	LSTM	0.93	0.89	0.90	0.91
	CLSTM	0.96	0.97	0.96	0.96
7 Class	CNN	0.91	0.90	0.90	0.92
	LSTM	0.88	0.92	0.90	0.90
	CLSTM	0.89	0.95	0.92	0.93
9 Class	CNN	0.92	0.88	0.90	0.91
	LSTM	0.88	0.86	0.87	0.86
	CLSTM	0.93	0.92	0.92	0.92

the time series features well; it hence represents poor performance in many class scenarios. CLSTM shows high accuracy in small classes, but the accuracy decreases dramatically as the number of classes increases. This is because the data size is small, like the result of LSTM.

Our proposed data augmentation techniques expands the temporal information of the original data set. It is a proper for time series data. CNN with data augmentation improves the accuracy slightly (i.e., approximately 2%). However, in LSTM and CLSTM, the accuracy increase is very large (i.e., approximately 9%). This is a result of the expansion of the temporal information of the input data set. LSTM still shows lower accuracy than CNN because it cannot distinguish spatial features, but CLSTM shows much higher accuracy than CNN.

Fig. 9 depicts the confusion matrix of the seven- to nine-class classification. The recall rate of User 7 in Fig. 9(a) is 100%, but the precision is low at 71%. The 100% recall means that the model properly determined that all data of User 7 belonged to User 7. On the other hand, the accuracy of 71% means that the model incorrectly judged that the data of other people were those of User 7 29% of the time. In the case of User 4, unlike User 7, the recall rate was 77% and the precision was 100%. Because the model incorrectly judged that the 7 data entries of User 4 belonged to User 7, the recall rate was as low as 77%. Fig. 9 (b) presents a similar confusion matrix for the seven-class classification. Here, the data of User 4 and User 5 are ascribed to User 7, and the model repeatedly judged wrongly. Table 3 shows the evaluation parameters for CNN, LSTM, and CLSTM networks with and without data augmentation.

VI. CONCLUSION

This paper proposed a human identification system using Wi-Fi CSI data. It enables the reduction of data collection costs via data augmentation techniques such as a sliding window and time warping. These augmented data is input into deep learning models improving accuracy and enabling identification of more people. We evaluated the performance using three representative deep learning models, and the CLSTM suitable for dealing with time series data showed the highest accuracy and stable loss. When using the CLSTM network, it showed an accuracy of over 90% in all classes. The proposed system had the advantages of improving accuracy, increasing the number of subjects, and reducing the cost of data collection. However, it still had some limitations which need to be overcome for it to be used practically. If more than one subject walked on the installed system, the human identification system did not work properly. In the future, we plan to study data preprocessing techniques that classify mixed data to enable one more subject identification.

REFERENCES

[1] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.

[2] S.-R. Lee, S.-D. Joo, and C.-W. Lee, "An enhanced dynamic framed slotted ALOHA algorithm for RFID tag identification," in *Proc. Int. Conf. Mobile Ubiquitous Syst., Netw. Services (MobiQuitous)*, Jul. 2005, pp. 166–172.

[3] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Efficient iris recognition by characterizing key local variations," *IEEE Trans. Image Process.*, vol. 13, no. 6, pp. 739–750, Jun. 2004.

[4] R. Cappelli, M. Ferrara, and D. Maltoni, "Minutia cylinder-code: A new representation and matching technique for fingerprint recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2128–2141, Dec. 2010.

[5] K. Kasantikul, C. Xiu, D. Yang, and M. Yang, "An enhanced technique for indoor navigation system based on WIFI-RSSI," in *Proc. 7th Int. Conf. Ubiquitous Cloud Comput. (ICUFN)*, Jul. 2015, pp. 513–518.

[6] S. Sigg, U. Blanke, and G. Tröster, "The telepathic phone: Frictionless activity recognition from WiFi-RSSI," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun. (PerCom)*, Mar. 2014, pp. 148–155.

[7] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. ACM MobiCom*, Sep. 2015, pp. 65–76.

[8] Y. Ma, G. Zhou, and S. Wang, "WiFi sensing with channel state information: A survey," *ACM Comput. Surv.*, vol. 52, no. 3, pp. 1–36, Jun. 2019.

[9] T. Um, F. M. J. Pfister, D. Pichler, S. Endo, M. Lang, S. Hirche, U. Fietzek, and D. Kulić, "Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks," in *Proc. 19th ACM Int. Conf. Multimodal Interact. (ICMI)*, Nov. 2017, pp. 216–220.

[10] X. Cui, V. Goel, and B. Kingsbury, "Data augmentation for deep neural network acoustic modeling," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, no. 9, pp. 1469–1477, Sep. 2015.

[11] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, no. 2, pp. 1–32, Nov. 2013.

[12] W. Xi, J. Zhao, X.-Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang, "Electronic frog eye: Counting crowd using WiFi," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2014, pp. 361–369.

[13] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2015, pp. 65–76.

[14] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, Apr. 2016, pp. 1–12.

[15] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "WiFi-ID: Human identification using WiFi signal," in *Proc. Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, May 2016, pp. 75–82.

[16] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776, Jan. 2017.

[17] M. N. A. Nipu, S. Talukder, M. S. Islam, and A. Chakrabarty, "Human identification using WiFi signal," in *Proc. Joint 7th Int. Conf. Informat., Electron. Vis. (ICIEV), 2nd Int. Conf. Imag., Vis. Pattern Recognit. (icIVPR)*, Jun. 2018, pp. 1–2.

[18] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11n traces with channel state information," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, p. 53, Jan. 2011.

[19] Y. Zhang and H. Liu, "Impact of time-selective fading on the performance of quasi-orthogonal space-time-coded OFDM systems," *IEEE Trans. Commun.*, vol. 54, no. 2, pp. 251–260, Feb. 2006.

[20] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity Wi-Fi," *IEEE Trans. Mobile Comput.*, vol. 18, no. 6, pp. 1342–1355, Jun. 2019.

[21] Y. Jin, W.-S. Soh, and W.-C. Wong, "Indoor localization with channel impulse response based fingerprint and nonparametric regression," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 1120–1127, Mar. 2010.

[22] J. Zhang, F. Wu, B. Wei, Q. Zhang, H. Huang, S. W. Shah, and J. Cheng, "Data augmentation and dense-LSTM for human activity recognition using WiFi signal," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4628–4641, Mar. 2021.

[23] N. Jaitly and G. E. Hinton, "Vocal tract length perturbation (VTLP) improves speech recognition," in *Proc. ICML Workshop Deep Learn. Audio Speech Lang*, vol. 117, 2013, pp. 1106–1114.

[24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst*, May 2017, pp. 1097–1105.

[25] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016.

[26] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "WFID: Passive device-free human identification using WiFi signal," in *Proc. 13th Int. Conf. Mobile Ubiquitous Syst., Comput., Netw. Services (MOBIQUITOUS)*, Nov. 2016, pp. 47–56.

[27] J. Lv, W. Yang, D. Man, X. Du, M. Yu, and M. Guizani, "Wii: Device-free passive identity identification via WiFi signals," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2017, pp. 1–6.

HYUNGGEUN MO received the B.S. and M.S. degrees from the School of Electronics Engineering, Chungbuk National University, Cheongju-si, South Korea, in 2019 and 2021, respectively. His research interests include deep learning-based IoT applications and mobile computing.

SEUNGKU KIM received the B.S., M.S., and Ph.D. degrees from the School of Electrical Engineering, Korea University, Seoul, South Korea, in 2007, 2010, and 2013, respectively. He joined the Software Center Division, Samsung Electronics, South Korea, in 2013. Since September 2015, he has been an Assistant Professor at the School of Electronics Engineering, Chungbuk National University, South Korea. His research interests include mobile computing, wireless sensor networks, body area networks, mobile *ad hoc* networks, Bluetooth, and the Internet of Things.