A Lightweight Deep Learning Algorithm for Identity Recognition

Yanjie Cao[1], Zhiyi Zhou[1], Pengsong Duan[1(✉)], Chao Wang[1], and Xianfu Chen[2]

[1] School of Software, Zhengzhou University, Zhengzhou, China
{caoyj,duanps}@zzu.edu.cn, zhou_zhi_yi@163.com, austin423@126.com
[2] VTT Technical Research Centre of Finland, Espoo, Finland
xianfu.chen@vtt.fi

Abstract. The challenges in current WiFi based gait recognition models, such as the limited classification ability, high storage cost, long training time and restricted deployment on hardware platforms, motivate us to propose a lightweight gait recognition system, which is named as B-Net. By reconstructing original data into a frequency energy graph, B-Net extracts the spatial features of different carriers. Moreover, a Balloon mechanism based on the concept of channel information integration is designed to reduce the storage cost, training time and so on. The key benefit of the Balloon mechanism is to realize the compression of model scale and relieve the gradient disappearance to some extent. Experimental results show that B-Net has less parameters and training time and is with higher accuracy and better robustness, compared with the previous gait recognition models.

Keywords: Identity recognition · WiFi · Channel state information · Deep learning

1 Introduction

In recent years, identity recognition has been widely researched [8,10]. Different from the traditional biometrics used for identity recognition, such as fingerprint and iris, gait has attracted extensive attention due to the unique characteristics of long distance, non-contact, and not easy to disguise and so on [3]. Cameras [14], acceleration sensors [19], ground sensors [1] and other devices are used for collecting gait data. Nevertheless, these sensors need to be either highly sensitive and accurately located or worn by the monitored target. In contrast, a device-free approach is often more economical and convenient. However, such an approach may invade people privacy and is often susceptible to illumination. The authors in proposed to use Radio Frequency (RF) technology (such as the Doppler radar). But the dependence on a specialized hardware limits the deployment. On the other hand, WiFi is a promising alternative which utilizes the existing ubiquitous inhouse WiFi signal, frees people from any potable device, as well as avoids the influence of illumination and unnecessary personal

S. W. Loke et al. (Eds.): MONAMI 2020, LNICST 338, pp. 3–18, 2020.
https://doi.org/10.1007/978-3-030-64002-6_1

privacy invasion [4,6]. As a result, researchers begin to focus on WiFi-based identification. Shi et al. conducted a comprehensive analysis on the disturbance characteristics of CSI data based on daily behaviors such as walking and quiescence of human body, and realized the identity recognition function [5,12]. In the WiFi environment, Wang et al. realized user identification function by collecting and analyzing the disturbance effect of gait on CSI data [13]. Zeng et al. proposed an identification method, WiWho, which used multipath elimination and bandpass filtering to remove noise, and extracted gait characteristics from WiFi CSI data for identification, with an accuracy rate of 92% to 80% in a group of 2 to 6 people [17]. In the WiFi-ID method proposed by Zhang et al., disturbance characteristics of the spectrum of WiFi CSI data collected by human gait were utilized to achieve an identification rate of 93% to 77% in the case of 2–6 people [18]. Xin et al. proposed a gait based identification method, FreeSense, which processed the collected CSI data by principal component analysis, discrete wavelet transform and dynamic time normalization, respectively, achieving an identification rate of 94.5% to 88.9% in the case of 2 to 6 people [15]. The above work has achieved certain effects in WiFi based identification research. However, CSI describes the combined effects, caused by scattering, fading, doppler frequency shift and power attenuation, of transmitted signals [11]. Hence, traditional manual feature extraction method requires plenty of data preprocessing, whereas the extracted features are insufficiently effective to improve identification accuracy. Since CSI can characterize multipath propagation at subcarrier level, when behavior occurs, the CSI changes of subcarriers of different antenna pairs are related [2]. Therefore, we uses deep learning to extract spatiotemporal characteristics of CSI data and proposes in this paper a simple but effective gait recognition framework, B-Net, which has the potential of realizing automatic identity recognition. Compared with traditional manual feature extraction methods, B-Net extracts gait characteristics by taking the frequency energy graph as the input, which is converted from the original data. In particular, a Balloon mechanism is designed for autonomous gait feature extraction. We highlight the advantages of a B-Net: first, it gets rid of the tedious and meticulous data preprocessing such as denoising and artificial feature extraction; and second, it implements zeroburden identification, requiring no personal devices.

The main contributions of this paper are summarized as follows.

- We propose that the original CSI data from WiFi signal can be reconstructed and used for individual gait recognition.
- We propose a Balloon mechanism to make the network model of B-Net consume less storage and operation resources.
- We establish a data set based on WiFi signal for gait identification. WiFi devices collect gait data from volunteers of different genders, ages, heights and weights.
- In order to evaluate the performance of B-Net, we conduct extensive experiments under various parameter settings. Experimental results show that the system achieves improved performance. Particularly, the identification

accuracy of the system with a group of 50 people reaches 98.8%, which significantly reduces the number of participants and training time compared with conventional models.

The rest of this paper is organized as follows. In Sect. 2, we detail the B-Net architecture and analyze the CSI characteristics. We present the experimental results and discuss the open problems in Sect. 3. We finally draw conclusions in Sect. 4.

2 B-Net Architecture

In this section, we first overview the B-Net and then introduce the main modules.

2.1 Overview of B-Net

The structure-Net architecture is shown in Fig. 1, including data processing module, Deep feature extraction module and classification module. Data acquisition module: the system sets up a WiFi monitoring environment with a single emitter and a single receiver, which continuously collect the CSI data from different volunteers walking in the area. Data preprocessing module: it is responsible for preprocessing the obtained original CSI data by converting the data into frequency energy diagram, which is applicable to the convolutional layer. Balloon block: as the core algorithm of B-Net, the Balloon block derives the relation model of identity and CSI amplitude fluctuation through a compressed model with a small number of parameters, and extracts effective features from the preprocessed data. Classification module: according to different frequency energy graphs corresponding to different volunteers, the neural network is trained.

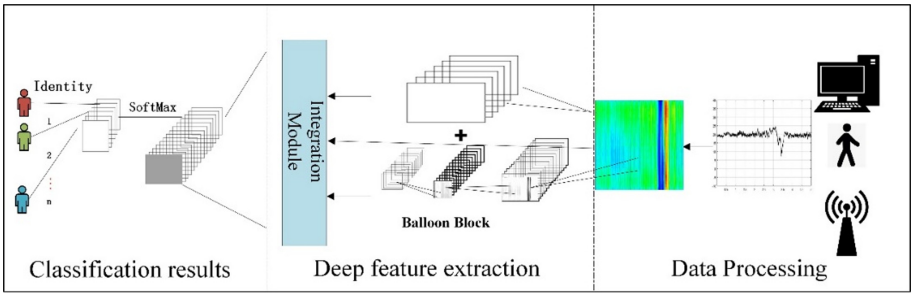


Fig. 1. Overview of B-Net.

2.2 Data Analysis Aonnd Representati

We assume that $X(f,t)$ and $Y(f,t)$ are, respectively, the frequency domain representations of transmitting and receiving signals at the WiFi carrier frequency f, that is,

$$Y(f,t) = H(f,t)X(f,t) \tag{1}$$

where

$$H(f,t) = \sum a_n e^{-j2\pi f\tau_n} + \sum a_m(t)\xi(f)e^{-j2\pi f\tau_m(t)} \tag{2}$$

with f being the carrier Frequency and $H(f,t)$ being the Channel Frequency Response (CFR) of the carrier. Due to the multipath effect, CFR can be represented as in Eq. 2, where $\sum a_n e^{-j2\pi f\tau_n}$ is the channel frequency response under static environment and $\sum a_m(t)\xi(f)e^{-j2\pi f\tau_m(t)}$ is the dynamic CFR value changing with time t. The CSI value of the$k$th subcarrier is expressed by $H(f_k)$,

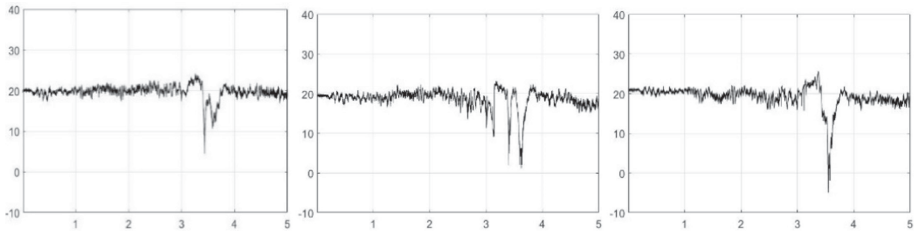
$$H(f_k) = \parallel H(f_k) \parallel e^{j\angle H(f_k)}, k \in [1,k] \tag{3}$$

where the amplitude and phase of the subcarrier are $\|H(f_k)\|$ and $\angle H(f_k)$, respectively. When collecting the CSI on subcarriers in WiFi signal, each CSI is a complex matrix of size $c*r*n$, where a fixed constant c represents the number of space streams, and r is the number of antennas on commercial routers. In this paper, the CFR value of a given antenna pair and a given OFDM subcarrier is called CSI stream, and hence there are $c*r*n$ CSI streams in the time series of CSI.

Current gait recognition studies usually organize CSI data into one-dimensional time series. The problem is that information in different subcarriers is redundant. Thus an optimal subcarrier selection algorithm is proposed in this paper to screen subcarrier signals with the best quality. Principal Component Analysis (PCA) is also applied to reduce the dimension of CSI data. The effects from different actions on signals are different, and the characteristics exhibit in not only the time dimension, but also the frequency dimension. Therefore, difference among the subcarriers is also one of the important features for action recognition.

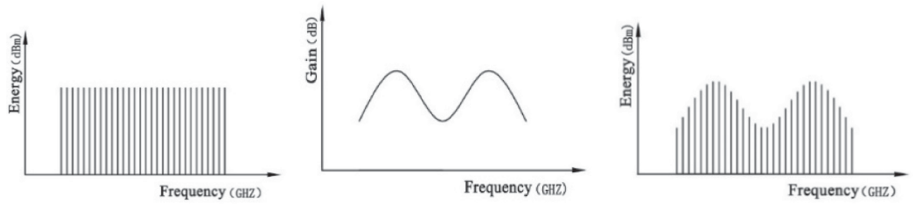
Figure 2 shows the influence of the same walking individual on different subcarriers. From experiments, it can be found that sensitivities of the first subcarrier, the 15th subcarrier and the 30th subcarrier are different. This phenomenon also exists in other subcarriers. The reason is that a subcarrier is sensitive to the movements of a particular part of the body, for example the arm or the leg, depending on whether the subcarrier wavelength and action size are comparable. Therefore, it is possible that different subcarriers contain different information when the same behavior occurs. As shown in Fig. 3, the energy distribution of subcarriers with different frequencies is evenly distributed during transmission, and the received energy distribution of each subcarrier varies significantly due to channel gain. Observing the similar distribution of different frequency of different pixel values in image, we propose a multidimensional CSI data organization

method, namely, the frequency energy graph, in which the length of time and the number of subcarriers are taken as length and width. Since the CSI received by different antennas to some extent reflects spatial position of the individual entering line-of-sight path, the number of receiving antennas is also taken as a dimension of frequency energy graph, which corresponds to color channel dimension of the ordinary RGB picture. After preprocessing original data at the receiving end, the corresponding frequency energy graph is constructed, whose format is similar to a RGB image.



(a) The 1^{st} subcarrier (b) The 15^{st} subcarrier (c) The 30^{st} subcarrier

Fig. 2. Sensitivities of different subcarriers to the same action.



(a) Emission (b) Channel gain (c) Receive

Fig. 3. Influence of wireless channel on subcarrier energy distribution.

Figure 4 shows the corresponding frequency energy graphs of CSI signals of volunteer A and volunteer B randomly selected from the data set in three different test cycles. By comparing the frequency energy diagram, it can be found that gait of the detected people shows unique individual characteristics on the frequency energy diagram. As can be seen from Fig. 4, there is a visible similarity between the three frequency energy graphs corresponding to the gait of one volunteer, whereas the frequency energy graphs corresponding to the gait of volunteer A and volunteer B have obvious discernibility. Compared with traditional data organization method, frequency energy graph contains both the time relationship on a single subcarrier and the interrelationship between different subcarriers. The first row are three groups of data randomly selected from

volunteer A data set, while the second row are three groups of data randomly selected from volunteer B data set.

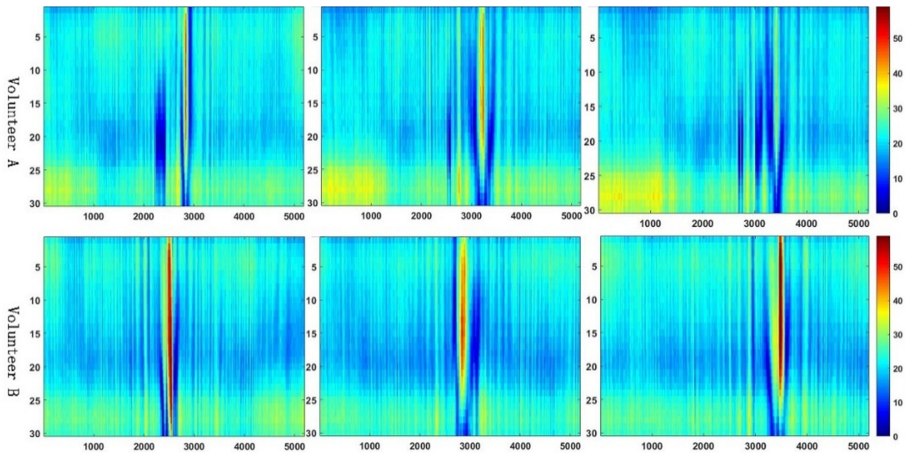


Fig. 4. The frequency energy diagram corresponding to different gait.

2.3 Network Architecture

Considering the timeliness and complexity of frequency energy graph obtained after processing the gait data, we propose a deep neural network model with a Balloon module. The function of each module is discussed as follows:

B-Net model combines one-dimensional convolutional kernel and three-dimensional convolutional kernel to realize feature extraction and channel information integration of data flow in the network. As being commonly known, increasing convolution kernel size increases the computation amount of the convolution operation. The larger the convolution kernel at the same sliding step, the higher the computation. From experiments, it is found that a larger convolution kernel receptive field size can be obtained by multiple small convolution kernel stacks. Using a small convolution kernel can obtain more abundant features than using a large convolution kernel, and make the decision function more discriminative. In addition, a three-dimensional convolution kernel is sufficient to capture the feature changes. The superposition of multilayer small convolutional kernel is equivalent to a single large convolutional kernel. However, with a deeper network, extra nonlinear ReLU function and diversity of features make the network larger in capacity and stronger in classification ability. The use of a large convolution kernel should be avoided from the perspective of model compression. How to use fewer parameters to get a deeper network is the objective of this paper. Overall, the model designed in this paper adopts the combination of a three-dimensional convolution kernel and a one-dimensional convolution kernel

to process the data. Note that the three-dimensional convolution kernel is mainly used to extract the features in frequency energy graph. The one-dimensional convolution kernel is shown in Fig. 5, where W and H are the length and width of the feature graph, and D is the number of feature graphs and channels. Although there is only one parameter, when this parameter is applied to a multichannel characteristic graph, it is equivalent to a linear combination of different channels. One-dimensional convolution learns the relationship between feature graphs and has the effect of decoupling, which reduces the number of parameters without compromising the model expressiveness.

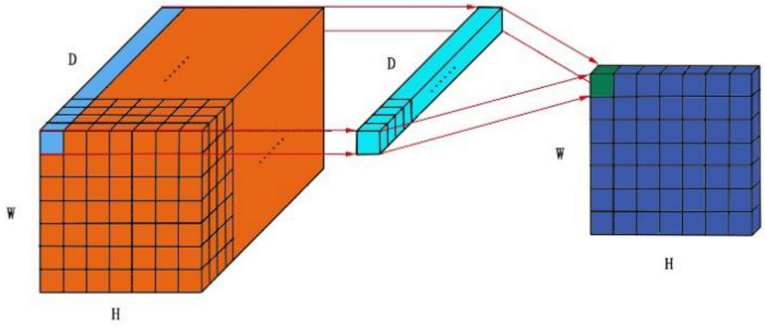


Fig. 5. One-dimensional convolution kernel diagram.

The one-dimensional convolution kernel is shown in Fig. 5, where W and H are the length and width of the feature graph, and D is the number of feature graphs and channels. Although there is only one parameter, when this parameter is applied to a multichannel characteristic graph, it is equivalent to a linear combination of different channels. One-dimensional convolution learns the relationship between feature graphs and has the effect of decoupling, which reduces the number of parameters without compromising the model expressiveness.

Nowadays, increasing the number of layers to improve accuracy has become the mainstream. However, in model training process, the number and size of convolution cores in the convolutional layer directly determine computational parameters of the layer and the amount of data input to the next layer. Blind stacking of network layers will result in an explosive increase in the computational amount. The primary goal of this paper is to build a CNN architecture with a simple structure, fewer parameters and comparable accuracy. In view of the above objectives, we design the Balloon mechanism. Figure 6 is a schematic diagram of Balloon mechanism. Among them, the review channel is inspired by the fact that human beings need to carry out visual input to the original graphics for many times when learning the knowledge of graphics, and a mechanism to simulate human beings' review behavior is designed. After the channel dimension of the original frequency domain, energy graph is increased by a one-dimensional convolution kernel, feature fusion is carried out with the output

feature graph of each layer. The fusion model can still achieve high classification accuracy on the basis of a small number of neurons. At the same time, because the feature map of each layer is added with the part of identity mapping, connection matrix degradation caused by layer deepening is reduced. After fusion, the feature map will complete crosschannel interaction and information integration through one-dimensional convolution kernel, so that the feature map can retrieve the original information from frequency energy map and reduce the dimension of feature map input to the next layer. Through review channel, each layer of convolutional network can access "low-level" features from the original data, so that the model can still guarantee good learning ability in the case of few convolutional cores. Classical convolutional neural networks usually increase the maximum pooling layer after convolutional layer to carry out downsampling, which is used to reduce the number of parameters and prevent model from overfitting. However, the characteristic density of frequency energy diagram in this paper is large, hence the classification accuracy is reduced by a pooling operation. Therefore, B-Net does not include a pool layer, but a regularization layer is added to fix the input of each neuron on the same distribution. Regularization layer speeds up the model training and uses a larger learning rate to train the network, hence reducing network retraining time. Finally, considering that the introduction of nonlinear factors can make the model more expressive, the linear rectifier function Relu is employed for activation.

2.4 Classification

B-Net aims to create a model that divides the gait type hypothesis into Y (1, 2. . . , k) kinds to predict the possible identity based on the input frequency energy diagram x. The purpose of B-Net model is to solve a multiclassification problem, so it is more appropriate to use the Softmax function, which maps the inputs to real numbers between 0 and 1, and normalizes the guaranteed sum to 1. The probability of each output category $\tilde{y}_i$ is calculated as

$$\tilde{y}_i = P(y|X) = \frac{\exp(z_y)}{\sum_{y=1}^{k} \exp(z_y)} y \in [1, k] \tag{4}$$

where z_y is the result of global pooling on the output. Global pooling operation can not only make the transformation between extracted feature map and final classification result more simple and natural, but also eliminate the need for a large number of training parameters, thus improving robustness and antioverfitting ability of the model. After obtaining $\tilde{y}_i$, Adam optimizer is used to minimize cross entropy loss between the predicted probability and the real one, which is given by

$$L = -\sum y_i \log(\tilde{y}_i) \tag{5}$$

3 Experiments and Evaluations

In this section, we use real gait data to verify the accuracy of the proposed lightweight neural network model through experiments.

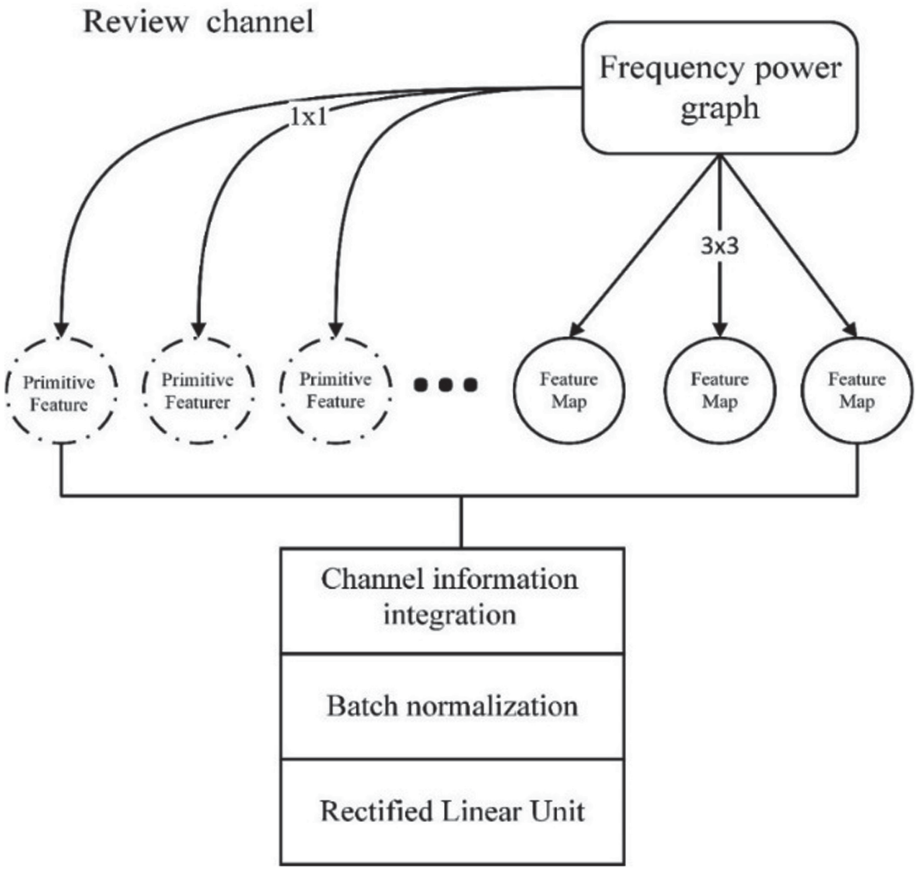
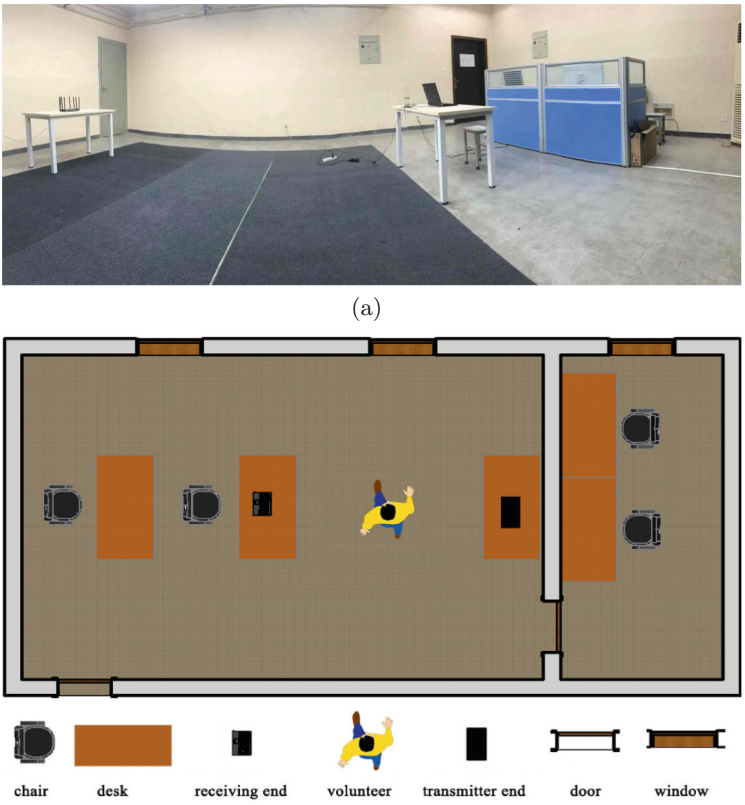


Fig. 6. Schematic diagram of Balloon mechanism.

3.1 Dataset Descriptions

In order to validate the performance from the model under different structures, we chose a relatively empty laboratory as shown in Fig. 7 for the experiments of multiperson gait recognition. Gait CSI data of 50 volunteers were collected as the training set and also as the test set. The volunteers, consisting of both male and female, ranged in ages from 18 to 30, ensuring the data diversity. During the data collection, each volunteer walked back and forth along the line-of-sight path between the vertical transmitter and the receiver for 5 s, and repeated the walking 80 times.



(a)

(b)

Fig. 7. Data acquisition environment.

3.2 Experiment Setup

In experiments, we used a TP_LINK AC1750 wireless router as the transmitter and a ThinkPad X201 portable computer terminal as the receiver. ThinkPad X201 is equipped with Intel 5300 802.11n WiFi NIC network card for receiving CSI data. The transmitter and receiver have 1 and 3 detectable antennas, respectively, forming 3 (1×3) CSI data streams. Each CSI data stream consists of 30 subcarriers, which are modulated by OFDM. Therefore, 90 $(1 \times 3 \times 30)$ CSI data streams were collected. In this paper, we carried out all experiments at the 5 GHz frequency to ensure data integrity and accuracy, and the sampling rate of CSI packets at the receiver was set to 1000 packets per second. The size of the sliding window was selected as $T = 600$ ms and the size of the sliding step was set as $d = 200$ to construct data samples and generate different frequency energy graphs. The overlap between the sliding window and the step size can ensure the continuity of data segmentation and help extracting feature information.

3.3 Key Performance Indicators

The multiperson gait recognition by the proposed B-Net system was evaluated by the accuracy, recall, precision and F_1 score. Specifically,

– Accuracy: the proportion of all correct predictions in the total number of predictions, which is given as

$$Accuracy = \frac{TP + TN}{TP + TN + FN + TN} \tag{6}$$

– Precision: the proportion of positive value of correct prediction to the total positive value of prediction, which defined by

$$Precision = \frac{TP}{TP + FN} \tag{7}$$

– Recall rate: the proportion of predicted positive value to the total actual positive value given as

$$Recall = \frac{TP}{TP + FP} \tag{8}$$

– F_1 score: the evaluation index of comprehensive recall rate and precision, which is used to comprehensively reflect the whole index,

$$F_1 = \frac{2TP}{2TP + FP + FN} \tag{9}$$

In above, TP (True Positive) represents the number of samples that are classified to be positive and actually positive ones, FP (False Positive) represents the number of samples that are classified to be positive but actually negative ones, FN (False Negative) represents the number of samples that are classified to be negative but actually positive ones, and TN (True Negative) represents the number of samples that are classified to be negative and actually negative ones.

3.4 Experimental Results

In order to explore the influence of different review channels on the experimental results, we designed two different review channels, namely, a Source link channel and a Fully link channel. A Source link channel represents that each network layer containing a Balloon mechanism accepts only the frequency energy graph raised by a one-dimensional convolution kernel. A Fully link channel represents that the nth network layer containing Balloon mechanism accepts n−1 feature graphs raised by a one-dimensional convolution kernel. As shown in Fig. 8, the walking data sets of 50 volunteers were used to train the neural networks with a Source link channel structure and a Fully link channel structure.

In Fig. 8, the horizontal axis represents the change in the number of network layers, while the vertical axis represents accuracy of the model in identifying data set. Meanwhile, the size of the model is marked at each marker in the

figure. It can be seen from the figure that both the Source link channel and the Fully link channel can maintain small model size. Moreover, with the increase of network layers, the accuracy of the Fully link channel case is better than that of the Source link channel case, but with a higher increasing rate. The purpose of this paper is to build a lightweight network. In other experiments, we used a Source link channel network.

Each layer of the Balloon mechanism in B-Net neural network carried out a crosschannel information integration processing after fusing the feature graph and reviewing the channel information. Information integration across channels is a kind of data dimensionality reduction in nature and has a certain decoupling function. At the same time, information integration processing can greatly reduce the number of training parameters, and hence an integration channel is the key to a lightweight model. To verify the effects of channel compression on model size and accuracy, we took a 3-layer Balloon network as an example, and the number of neurons in each layer of convolutional operation was set to 32, 64 and 128.

Table 1. Comparisons of experimental results of different compression channels.

Compression channel number	Model parameter number	Model size	Accuracy	Rate/model size
8/16/32	42260	0.6M	95	1.58
16/32/64	54154	0.748M	98	1.31
32/16/8	99298	1.3M	96.7	0.743
16/16/16	51522	0.71M	96.6	1.36
32/32/32	141282	1.8M	98	0.54
64/64/64	274562	3.4M	97.9	0.28

As shown in Table 1, channel compression number represents the number of feature graphs of the three-layer network after compression. This experiment compared the network with different compression degree and accuracy. The ratio of accuracy to model size is used as the key performance indicator for evaluating the model efficiency. It can be seen that each layer, as shown in the first and second rows of Table 1, got better results as the number of channels increased. The number of parameters of the model in the third row was much larger than that in the second row, but achieved lower accuracy. The reason behind is that the feature levels extracted by the convolution operation of each layer are different. The shallow convolution operation extracts "low-level" features of the data, which loses less information after dimensionality reduction. Deep convolution extracts more special and complex features, and dimension reduction of these features destroys the feature structure and causes the information loss. Results in Rows 4 to 6 show that the increase of parameters and model accuracy

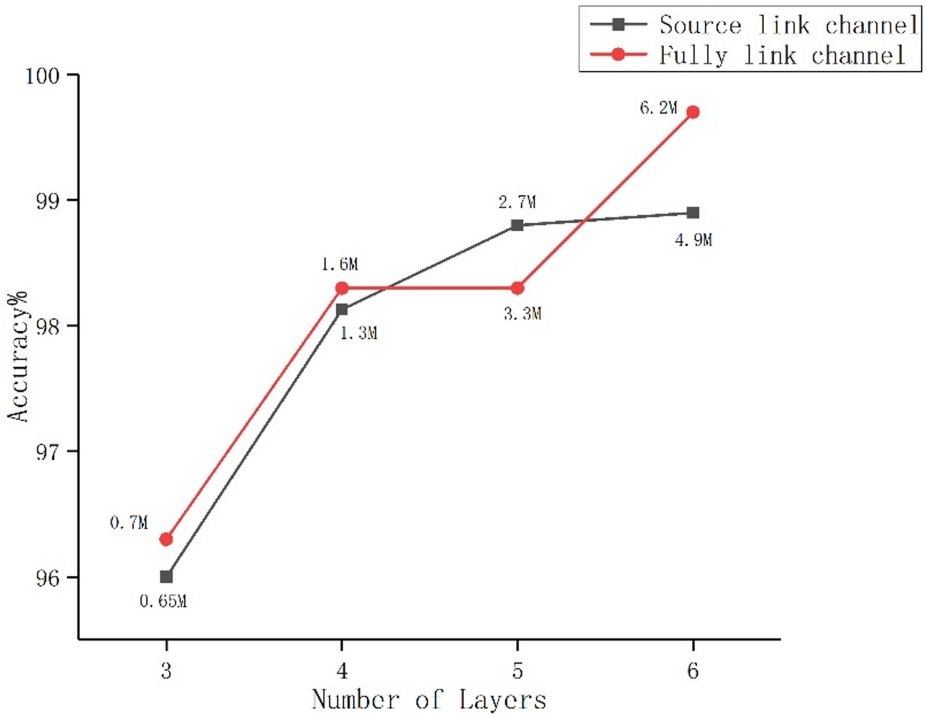


Fig. 8. Comparison of different channel structures.

consumed a large amount of storage space, and accuracy degrades when the number of parameters increased to a certain extent.

This experiment tells that the accuracy of the model relates to the parameter number. Increasing the number of parameters to improve model accuracy results in the storage consumption. Layer by layer compression using the Balloon mechanism saves model storage space while ensuring the accuracy. In addition to the comparison with other models, the experiment of gait data set of 50 people was taken as an example in the hall environment, using a GRU cyclic neural network [7], a full convolutional neural network FCN [9], a WiID model [16] and a Source link channel B-Net containing five layers of Balloon layer. In the experiment, GRU contained two GRU layers, FCN was composed of three convolutional layers, and the setting of WiID followed. All the above models showed good ability in the identification of 10 15 people from the data set. As can be seen from Fig. 9, the recognition accuracy, precision, recall rate and F1 score of GRU and WiID models all decreased significantly as the number of identification increased to 50. Although the evaluation criteria of FCN model are less attenuated, it can be seen that the model size is much larger than B-Net. Table 2 illustrates that B-Net can extract gait characteristics contained in CSI information more completely and accurately by using a frequency energy graph

and a Balloon mechanism, and consumes less storage resources. Its recognition performance was the best among the four, reaching 98.8% in accuracy, and the best in the comparison experiment of gait recognition based on Wi-Fi sensing.

Table 2. Experimental results of different models with 50 samples.

Model	Accuracy%	Recall rate%	Precision%	F_1 score	Model size
GRU	75.7	75.6	76	0.758	3.12M
WiID	65.6	65.4	66.0	0.657	2.2M
FCN	93.0	93.2	94.1	0.936	14M
B-Net	98.8	98.8	98.9	0.988	2.7M

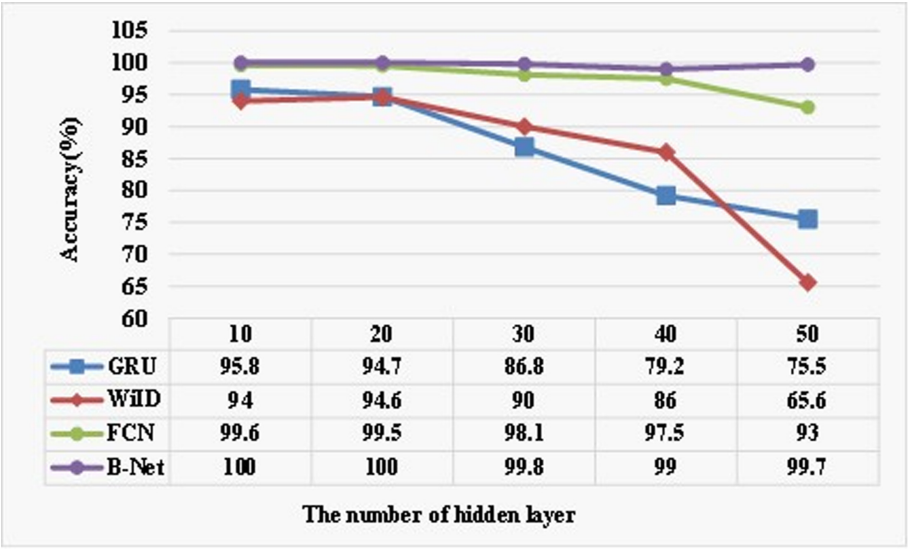


Fig. 9. Comparison of different models.

4 Conclusion

We propose a simple and efficient neural network model, named as B-Net, which can be used to identify WiFi signals from different individuals. Facing the challenges of long training time, large computation and poor performance of traditional deep neural networks, we design a compact neural network to identify CSI of different individuals. The proposed model can not only simplify the training complexity but also ensure good training performance. Compared with the state-of-the-art deep learning methods for gait recognition, the proposed approach has higher recognition rate and operational efficiency. Additionally, B-Net gets rid

of the dependence of hand-coded features and realizes end-to-end identity recognition. Further work will be based on B-Net and combine the idea of migration learning to investigate a generalized model for different indoor environments.

References

1. Al-Naimi, I., Wong, C., Moore, P., Chen, X.: Multimodal approach for non-tagged indoor identification and tracking using smart floor and pyroelectric infrared sensors. Int. J. Comput. Sci. Eng. **14**(1), 1–15 (2017)
2. Ali, K., Liu, A.X., Wang, W., Shahzad, M.: Keystroke recognition using WiFi signals. In: The 21st Annual International Conference on Mobile Computing and Networking, pp. 90–102 (2015)
3. Connor, P., Ross, A.: Biometric recognition by gait: a survey of modalities and features. Comput. Vis. Image Underst. **167**, 1–27 (2018)
4. Gu, Y., et al.: EmoSense: computational intelligence driven emotion sensing via wireless channel data. IEEE Trans. Emerg. Top. Comput. Intell. **4**(3), 216–226 (2019)
5. Gu, Y., Wang, Y., Liu, Z., Liu, J., Li, J.: SleepGuardian: an RF based healthcare system guarding your sleep from afar. IEEE Netw. **34**(2), 164–171 (2019)
6. Gu, Y., Zhang, X., Liu, Z., Ren, F.: BeSense: leveraging WiFi channel data and computational intelligence for behavior analysis. IEEE Comput. Intell. Mag. **14**(4), 31–41 (2019)
7. Kyunghyun, C., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv:1406.1078 (2014)
8. Lin, N., et al.: Contactless body movement recognition during sleeping via WiFi signal. IEEE Internet Things **7**(3), 2028–2037 (2019)
9. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)
10. Mastali, N., Agbinya, J.I.: Authentication of subjects and devices using biometrics and identity management systems for persuasive mobile computing: a survey paper. In: 2010 Fifth International Conference on Broadband and Biomedical Communications (IB2Com) (2010)
11. Ohara, K., Maekawa, T., Matsushita, Y.: Detecting state changes of indoor everyday objects using Wi-Fi channel state information. Proc. ACM Interact. Mob. Wearable Ubiquit. Technol. **1**(3), 1–28 (2017)
12. Shi, C., Liu, J., Liu, H., Chen, Y.: Smart user authentication through actuation of daily activities leveraging WiFi-enabled IoT. In: The 18th ACM International Symposium, pp. 1–10 (2017)
13. Wang, W., Liu, A.X., Shahzad, M.: Gait recognition using WiFi signals. In: ACM International Joint Conference Pervasive Ubiquitous Computing, pp. 363–373 (2016)
14. Wu, Z., Huang, Y., Wang, L., Wang, X., Tan, T.: A comprehensive study on cross-view gait based human identification with deep CNNs. IEEE Trans. Pattern Anal. Mach. Intell. **39**(2), 209–226 (2017)
15. Xin, T., Guo, B., Wang, Z., Li, M., Yu, Z., Zhou, X.: Freesense: indoor human identification with WiFi signals. In: IEEE Global Communications Conference (GLOBECOM), pp. 1–7 (2016)

16. Yu, X., Chen, W., Wand, D.: A deep learning algorithm for contactless human identification. J. Xi'an Jiaotong Univ. **53**(04), 128–133 (2019)
17. Zeng, Y., Pathak, P.H., Mohapatra, P.: WiWho: WiFi-based person identification in smart spaces. In: 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), pp. 1–12 (2016)
18. Zhang, J., Wei, B., Hu, W., Kanhere, S.S.: WiFi-ID: human identification using WiFi signal. In: 2016 International Conference on Distributed Computing in Sensor Systems (DCOSS), pp. 75–82 (2016)
19. Zhang, Y., Pan, G., Jia, K., Lu, M., Wang, Y., Wu, Z.: Accelerometer-based gait recognition by sparse representation of signature points with clusters. IEEE Trans. Cybern. **45**(9), 1864–1875 (2015)