Journal of Computer and Communications, 2023, 11, 15-28
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Human Activity Recognition Based on Frequency-Modulated Continuous Wave and DenseNet

Wenshuo Jiang, Yuqian Ma, Wencheng Zhuang, Zhongqiang Wu, Yiming Hua, Meng Li, Zhengjie Wang*

College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao, China
Email: w1360570997@163.com, 2206911173@qq.com, 2831599729@qq.com, wuzq917@163.com, 1982911606@qq.com, 1730841916@qq.com, *cieewangzj@163.com

How to cite this paper: Jiang, W.S., Ma, Y.Q., Zhuang, W.C., Wu, Z.Q., Hua, Y.M., Li, M. and Wang, Z.J. (2023) Human Activity Recognition Based on Frequency-Modulated Continuous Wave and DenseNet. *Journal of Computer and Communications*, **11**, 15-28.
https://doi.org/10.4236/jcc.2023.117002

Received: May 27, 2023
Accepted: July 11, 2023
Published: July 14, 2023

Abstract

With the development of wireless technology, Frequency-Modulated Continuous Wave (FMCW) radar has increased sensing capability and can be used to recognize human activity. These applications have gained widespread attention and become a hot research area. FMCW signals reflected by target activity can be collected, and human activity can be recognized based on the measurements. This paper focused on human activity recognition based on FMCW and DenseNet. We collected point clouds from FMCW and analyzed them to recognize human activity because different activities could lead to unique point cloud features. We built and trained the neural network to implement human activities using a FMCW signal. Firstly, this paper presented recent reviews about human activity recognition using wireless signals. Then, it introduced the basic concepts of FMCW radar and described the fundamental principles of the system using FMCW radar. We also provided the system framework, experiment scenario, and DenseNet neural network structure. Finally, we presented the experimental results and analyzed the accuracy of different neural network models. The system achieved recognition accuracy of 100 percent for five activities using the DenseNet. We concluded the paper by discussing the current issues and future research directions.

Keywords

Human Behavior Recognition, Millimeter-Wave Radar, Convolutional Neural Networks, Wireless Signal

1. Introduction

With the continuous development of wireless technology and the increasing demand for human-computer interaction, using wireless signals to recognize human behavior has become a popular research topic [1]. Recognizable behaviors mainly include daily activity recognition, crowd counting, vital sign detection, fall detection, identity authentication, gesture recognition, and more [2] [3]. The wireless signals used for behavior recognition mainly include WiFi signals and radar signals. WiFi devices are widely deployed, and WiFi signals are easy to obtain, making them highly practical [4]. Radar signals can provide high resolution and throughput, enabling fine-grained human behavior recognition [5]. These two typical wireless signals have been widely used in human activity recognition and have achieved significant results. In the following section, we will review these two common behavior recognition techniques.

1) Human activity recognition based on WiFi signals

With the widespread deployment of WiFi devices, using WiFi signals for behavior recognition has gained widespread interest. In 2011, Halperin *et al.* modified the network card driver of WiFi devices and released the CSI Tool script to extract channel state information [6] from commercial wireless network devices. The channel state information can be obtained from WiFi devices and can provide finer granularity, leading to widespread research.

CARM [7] is a typical activity recognition system that analyzes the relationship between human motion speed and specific activities through the CSI-Speed module and the CSI-Activity module to realize daily activity recognition. E-Eyes [8] uses matching algorithms to recognize 11 kinds of *in-situ* activities and 8 kinds of non-*in-situ* activities through analysis of CSI amplitude information. WiFall [9] is a typical fall detection system that uses a sharp drop in signal frequency as a judgment indicator to detect falls. WiseFi [10] establishes a signal arrival angle model to analyze the relationship between the CSI amplitude phase and activity to achieve activity recognition.

We find that WiFi-based human activity recognition has made certain progress in many aspects, but also has some disadvantages, such as low resolution, weak signal interpretability, decreased accuracy in multi-person states, and more.

2) Human activity recognition based on FMCW signals

Although WiFi-based human activity recognition has made some progress, there are challenges due to the limited resolution of WiFi signals. When fine resolution is needed, FMCW is a more suitable choice. In addition to the advantages of traditional WiFi, FMCW provides more data, finer resolution, and more physical features. As a result, FMCW can offer more potential applications, including human action recognition, fall detection, and vital monitoring.

Human Activity Recognition (HAR) is one of the major applications of FMCW radar [11]. Ding *et al.* [12] proposed a novel Dynamic Range-Doppler Trajectory (DRDT) method using FMCW radar to identify continuous human motions in emulating real-life scenarios. Ding *et al.* [13] combined sparse theory and PointNet network and utilized both the Time-Doppler (TD) and Range-Doppler (RD) domains to recognize human motion.

Fall detection is another typical application of FMCW radar. Saeed *et al.* [14]

designed a recognition system to effectively identify falls/collapses and categorize other daily living activities, including sitting, standing, walking, drinking, and bending. Wang *et al.* [15] utilized a Line Kernel Convolutional Neural Network (LKCNN) to detect falling states, while Wang *et al.* [16] leveraged Pattern Contour-Confined Doppler-Time (PCC-DT) maps to recognize soft fall motions.

FMCW devices can also be used to monitor vital signals such as respiration and heart rate due to their fine resolution. For example, Alizadeh *et al.* [17] extracted the respiration and heart rates of a patient lying down on a bed. Turppa *et al.* [18] remotely monitored heart rate and respiration in normal and abnormal physiological conditions during sleep and obtained an average relative mean absolute error of 3.6% (86% correlation) and 9.1% (91% correlation) for heart rate and respiration rate, respectively. Sacco *et al.* [19] achieved high accuracy in both respiratory rate and heartbeat monitoring. ViMo [20] is a typical multiple-person vital sign monitoring system using commodity millimeter-wave radio and achieved a median error of 0.19 and 0.92 Breaths per Minute (BPM), respectively, for Respiration Rates (RRs) and Heart Rates (HRs) estimation.

In summary, FMCW radar has shown potential for various HAR applications, including remote health monitoring, indoor localization and tracking, and gesture recognition. The technology holds promise for future advancements in security, healthcare, and human-computer interaction.

Despite the notable performance achieved in the aforementioned studies, they typically construct specialized neural network models to achieve good recognition results. However, many typical deep learning models have been proven effective for feature extraction and classification in computer vision tasks. Therefore, using these neural networks for human activity recognition is a hot research topic. FMCW signals effectively contain human activity characteristics, so they can be employed to feed into typical neural networks to implement human activity recognition. This paper focuses on human activity recognition using FMCW. Specifically, the DenseNet model is used for human activity recognition using FMCW signals.

The contribution of this paper can be summarized as follows. We proposed a human activity system using the FMCW signal and the DenseNet technique. The system achieved 100% accuracy for five human activities conducted by five participants using the standard DenseNet model with adopted input and output shapes. This result shows that a typical neural network can be used for real human activity recognition using FMCW signals. It also indicates that some typical neural networks have a broad application scenario in human-computer interaction using FMCW signals.

2. Methodology

2.1. FMCW Signal

The frequency of a FMCW radar has a typical working characteristic where its signal changes linearly over time, creating a linear frequency-modulated pulse. The working principle of a linear frequency-modulated pulse signal in a radar system can be described as follows. Firstly, the signal source generates a linear frequency-modulated pulse and transmits it through the transmitting antenna (TX antenna). Then, the transmitted signal is reflected by the object to generate a reflected linear frequency-modulated pulse, which is received by the receiving antenna (RX antenna). Finally, the "mixer" mixes the TX-transmitted signal and the RX-received signal together, and a low-pass filter generates an Intermediate Frequency (IF) signal. This process is the processing of a linear frequency-modulated signal chirp. By performing the same processing on consecutive multiple chirps and then splicing them into a frame of data, a radar system can gather more information about the target object, such as its distance, velocity, and motion pattern.

2.2. Framework of the System

The fundamental principle of human activity recognition using FMCW can be depicted as follows. FMCW device transmits a continuous wave signal that is modulated in frequency. The transmitted signal is reflected back from the target object. When a person walks or conducts some activities within the range of FMCW signal coverage, the FMCW signal will undergo complex path changes and be reflected by the body. These changes can be used to recognize human activity by building the relationship between the signal variation and the activity categories. We can infer object properties or object activity by analyzing the reflected signal characteristics. This approach enables FMCW devices to recognize various human activities, such as walking, running, sitting, and standing, by analyzing the unique signal patterns produced by each activity [11].

The system framework for FMCW-based human activity recognition contains four parts: experimental data collection, data processing, and neural network training, and human activity recognition, as shown in **Figure 1**. The DJ-IWR1843

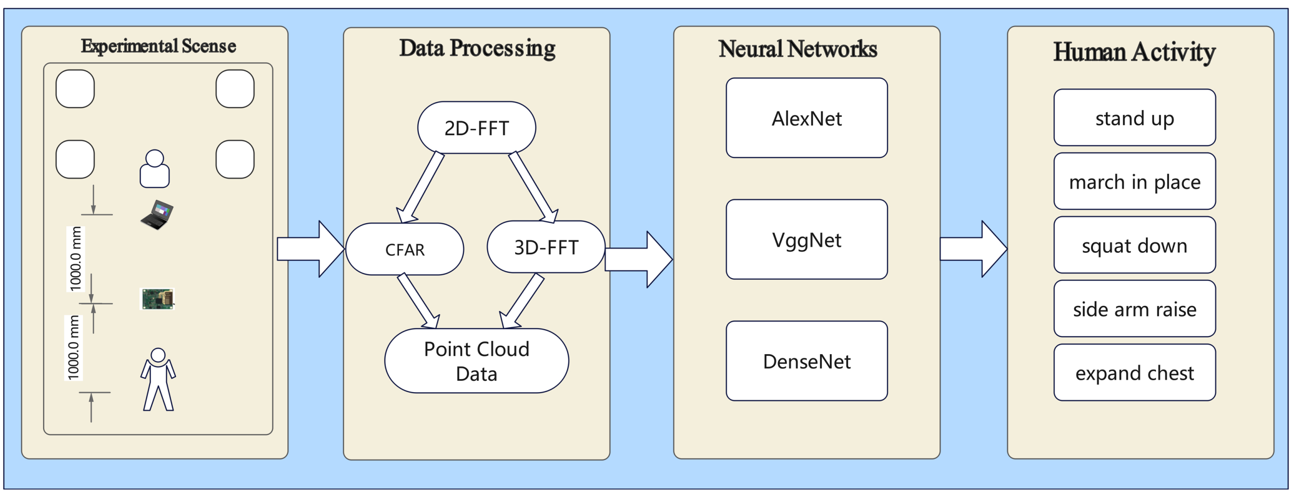

Figure 1. The framework of FMCW-based human activity recognition.

is used as the experimental device, with its basic parameters described in **Table 1**. The distance between the hardware and participants is set to one meter, and they perform their activities in place. The FMCW device is placed on a pole with a height of one meter above the ground.

During normal daily activities, the FMCW device generates a point cloud that can be collected. After collecting the point cloud data, the data are normalized and supplemented to feed them into a neural network. Finally, the preprocessed data are inputted into neural networks such as AlexNet [21], VggNet [22], and DenseNet [23] to evaluate the performance of the system. By analyzing the training and validation processing, human activity recognition is finally realized.

In this system, we define the activity as a set of human actions that can be recognized by the FMCW-based human activity recognition system. We have chosen five daily activities to evaluate the system performance, which are stand up, march in place, squat down, side arm raise, and expand chest. These activities are represented as activity A-E.

2.3. Neural Network Models

Deep learning is a type of machine learning algorithm that utilizes multiple layers of representation to learn complex patterns in large datasets [24]. It uses deep neural networks to complete complex tasks [25]. Deep neural networks typically consist of an input layer, multiple hidden layers, and an output layer. They automatically extract hidden features of the data using simple but non-linear layers, with each layer using the output of the previous layer as input [26]. Deep learning has made tremendous progress in advancing various artificial intelligence applications in recent years, particularly in fields like computer vision, automatic speech recognition, natural language processing, and medical image analysis, where it has achieved significant success [27].

In this paper, we utilize the typical DenseNet model to classify human activity using FMCW signals. Specifically, we employ the DenseNet121 model to categorize the activity using FMCW. Furthermore, we compare our model with other commonly used neural network models, such as AlexNet and VggNet, to

Table 1. The main parameters of DJ-IWR1843 device.

Radar working parameter	Value
Initial frequency (GHz)	76
Bandwidth (GHz)	4
Number of transmitting antennas	3
Number of receiving antennas	4
Frequency growth slope (MHz/μs)	60
Range resolution (cm)	3.75
Frame period (ms)	50
Frame rate (fps)	20

verify the effectiveness of our system. We described the structure of AlexNet, VggNet, and DenseNet according to the reference [28].

AlexNet is a Convolutional Neural Network (CNN) proposed in 2012 for image classification. It demonstrated the power of CNNs and GPU-based deep learning. AlexNet is comprised of 8 layers, consisting of 5 convolutional layers and 2 fully-connected layers, as depicted in **Figure 2**. It utilizes rectified linear units for nonlinearity, max pooling layers, dropout, and softmax loss. AlexNet achieved a top-5 error rate of 15.4% on ImageNet, far surpassing previous results. AlexNet's architecture and training procedures established a template for modern CNNs.

The VGG architecture is derived from the name of the study group (Visual Geometry Group) that proposed it in 2014. It is a convolutional neural network architecture that utilizes a simple design with only 3 × 3 convolutional layers

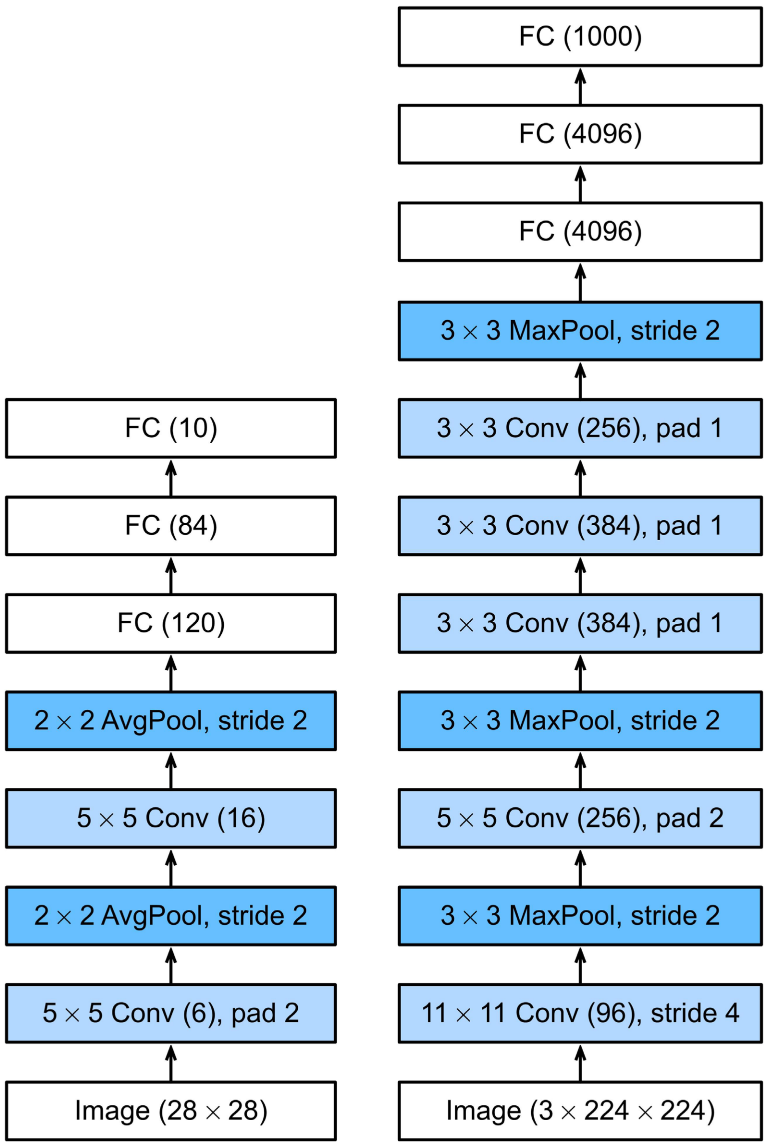

Figure 2. The basic structure of the AlexNet.

stacked on top of each other in increasing depth, as depicted in Figure 3. This small receptive field is conducive to capturing more complex features. VGG achieved state-of-the-art accuracy on ImageNet with 16 - 19 layers during the competition. Two major variants, VGG16 and VGG19, have 16 and 19 layers, respectively. VGG emphasizes the importance of model depth, and its simplicity and strong performance have made it an influential CNN architecture.

DenseNet is a convolutional neural network architecture that uses dense connections between layers. In this architecture, each layer obtains additional inputs from all preceding layers and passes on its own feature-maps to all subsequent

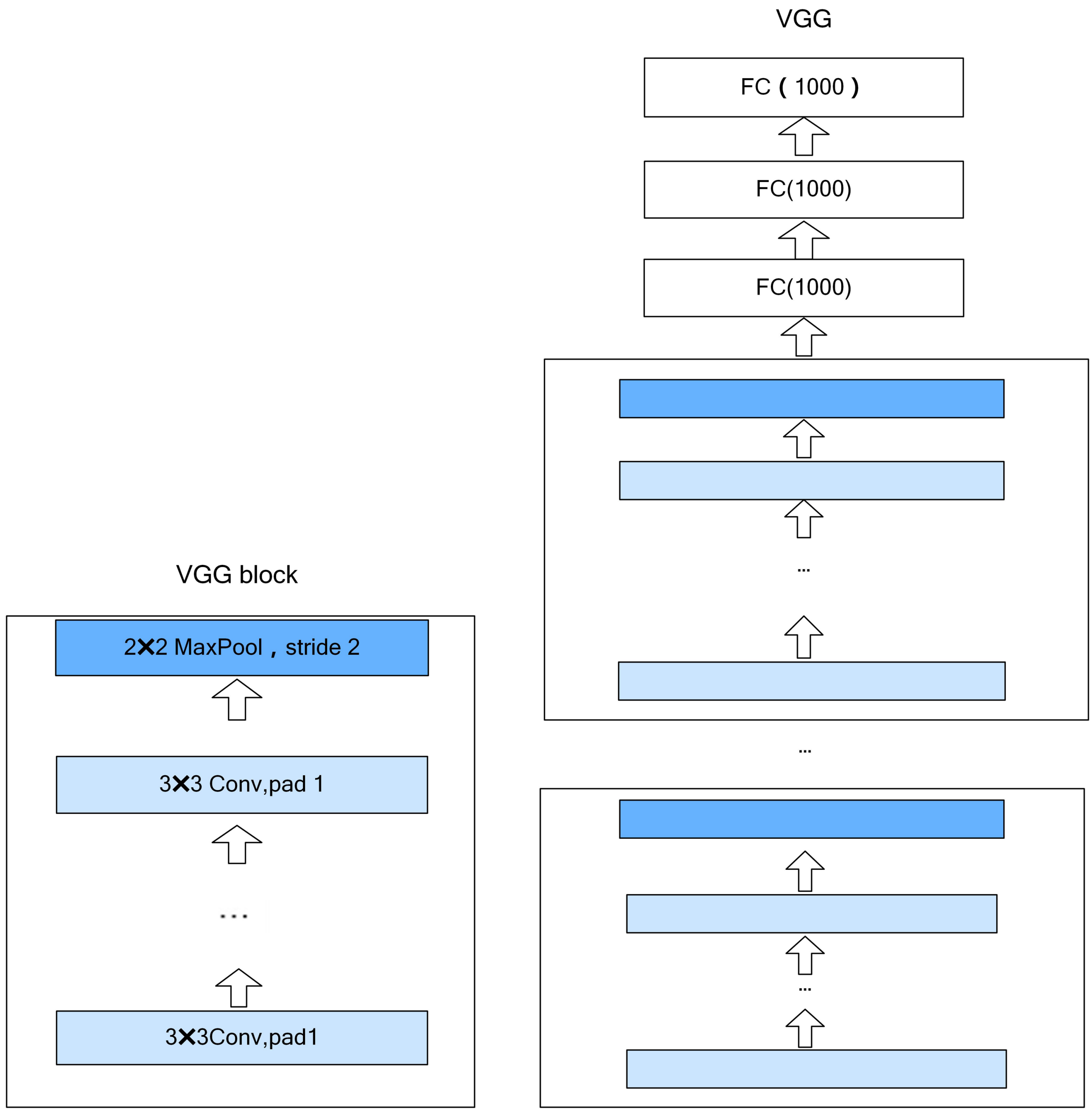

Figure 3. The basic structure of the VggNet.

layers, as depicted in **Figure 4**. The use of dense connections can alleviate the vanishing-gradient problem, strengthen feature propagation, reduce parameter redundancy, and improve learning efficiency. DenseNet models have achieved very good results on image classification and other tasks. DenseNet demonstrates the power of strengthened connections between network layers for more sophisticated learning.

The network parameters of the neural network adopted by this system are as follows. The system utilizes the PyTorch framework to implement the neural network and modifies the input channel of the first convolution and the last classification number. The input channel is set to six since the point cloud has six feature dimensions, and the classification number is changed to five since the system recognizes five activities.

2.4. Results

For the experiment, we invited 5 college students (two girls and three boys) to perform the activities. Specifically, each student collected about 50 samples for each activity, resulting in a total of approximately 2500 samples. The total samples were randomly split into training, validation, and test sets in a 7:2:1 ratio. We used the DenseNet neural network for the system and modified the network parameters according to our FMCW data.

1) The FMCW data processing

Each data frame that we collect contains six fields, which include the x, y, and z coordinates of the subjects, Signal-to-Noise Ratio (SNR), velocity, and noise. We found that different activities could result in a different number of point clouds, leading to a wide variation in the number of point clouds that we could collect. Because we use a neural network to implement activity classification, we need a fixed data shape to feed the data into the network. The hardware sends data at 20 frames per minute, and we measured the data and found that the maximum number of point clouds is less than 80 when the subject performs various activities. We spent three seconds to perform an activity. So, we obtained 60 frames data since the rate of frame is twenty. We set the shape of one activity frame as 60 × 80 × 6, where 60 is the number of frames during the activity time, 80 is the maximum point cloud count, and 6 is the dimension of one point. For frames that own less points, we supplement it with 0 to obtain a fixed shape for the activity.

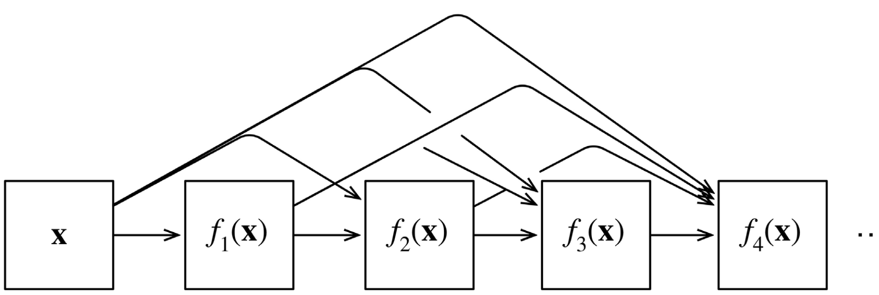

Figure 4. The basic structure of the DenseNet.

To ensure that the different units and values of the six fields in the data frame do not affect the computation process and recognition accuracy, we perform data normalization at each dimension. We normalize each field to have a mean of zero and a variance of one, which eliminates the adverse effects on recognition accuracy.

2) The neural network model design

In this paper, we use the DenseNet model to implement human activity recognition. We also compare the performance of the DenseNet model with two other typical models, AlexNet and VGG. Due to the differences between the original FMCW data and the image data that is fed into the neural network, we must slightly modify the DenseNet parameters.

Firstly, we modify the in_channel of conv0 of DenseNet features to 6 since we have six dimensions of FMCW point cloud. Secondly, we change the classifier of the DenseNet to have five categories.

To train the neural network, we utilize the cross-entropy loss function and the Adam optimization method. To enhance training effectiveness, we set the initial learning rate as 0.01 and dynamically adjust the learning rate by multiplying 0.1 when the validation accuracy plateaus. We initialize the original parameters of the neural network using the kaiming approach for each model. We can achieve more recognition results using a small epoch since we employ typical neural network. Therefore, we set the epoch as 100 and the batch size as 16.

3) The result of the experiment

Figure 5 displays the training and validation accuracies plotted against the epoch. Our findings indicate that due to the utilization of kaiming network

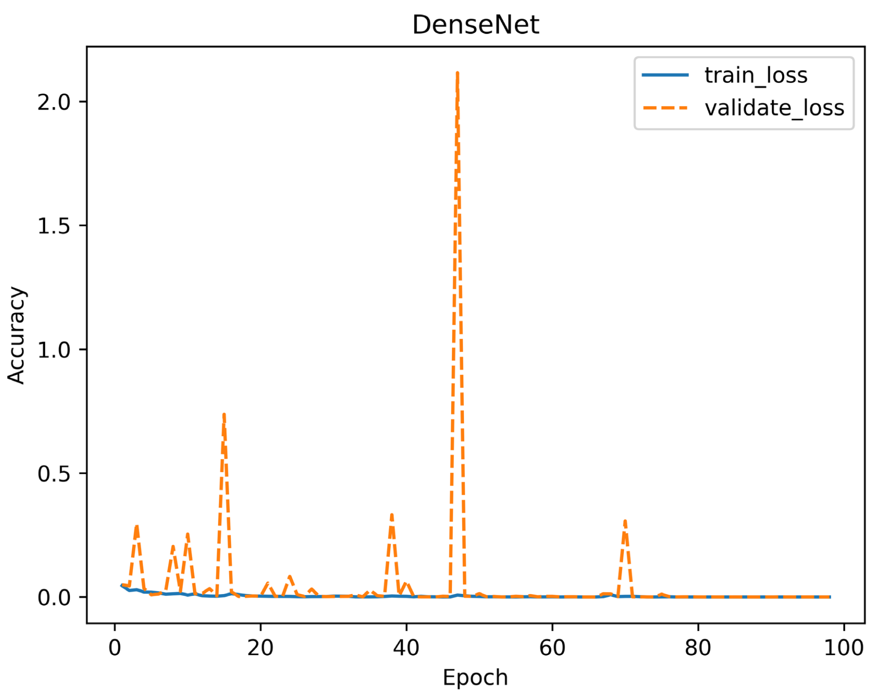

Figure 5. DenseNet training accuracy

parameter initialization, the training process converges rapidly after ten epochs.

In addition, we draw the confusion matrix to evaluate the activity recognition results, as shown in **Figure 6**. We observed that every each activity is always 100 percent on the test data, indicating that the activity recognition results are accurate and consistent.

2.5. Discussion

In this section, we compare the recognition accuracy of various neural network models, such as AlexNet and VggNet, which are commonly used in image recognition applications, with that of DenseNet. This comparison can effectively prove that DenseNet has great performance in recognizing activity using FMCW signal.

An analysis was first conducted regarding the influence of disparate training durations across three neural network architectures. It was ascertained that the DenseNet models reach convergence with celerity. At epoch 30 to 50 or after 70, the training loss and validation loss keep a similar trend and are almost equal. During other epochs, the training loss always decreases, and the validation loss has much variation. However, this variation may come from parameter updating and cannot affect test accuracy. We found the test accuracy kept at 100% after 30 epochs for DenseNet, even though the validation had some fluctuation. Different from DenseNet, the other two models have the same feature. The validation loss keeps at 0.1 after some initial fluctuations. This result shows that the training process does not converge, and the difference between training loss and validation loss always has about 0.05. The result shows that DenseNet has great feature

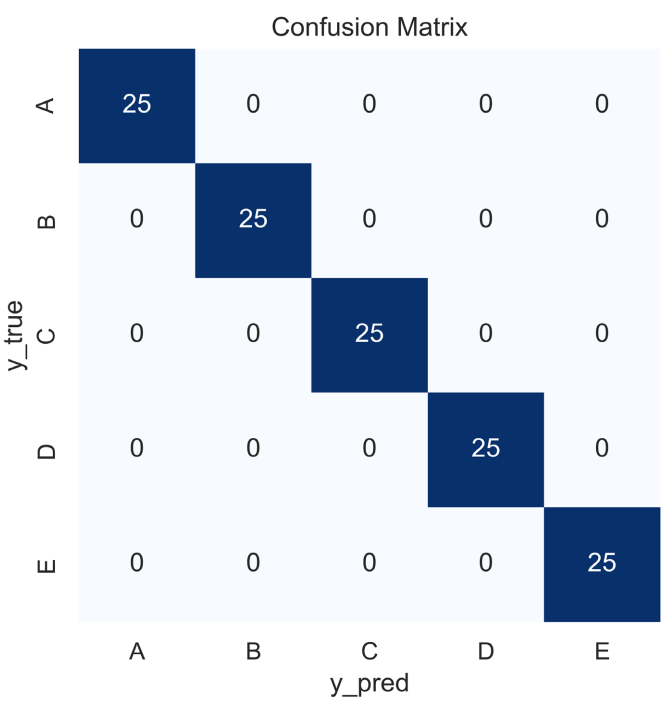

Figure 6. The confusion matrix of DenseNet for activity recognition.

extraction and activity recognition ability.

Next, we analyze the recognition accuracy among these three models. We can find that the recognition accuracy reaches 100% for five actions and five subjects. Therefore, DenseNet achieves the best test results for a neural network model. Differently, the other two models cannot implement activity classification since the training process does not converge, even though we dynamically reduce the learning rate, as shown in Figure 7. Besides, we find that almost all test samples have been classified into three activities for the other models. So,

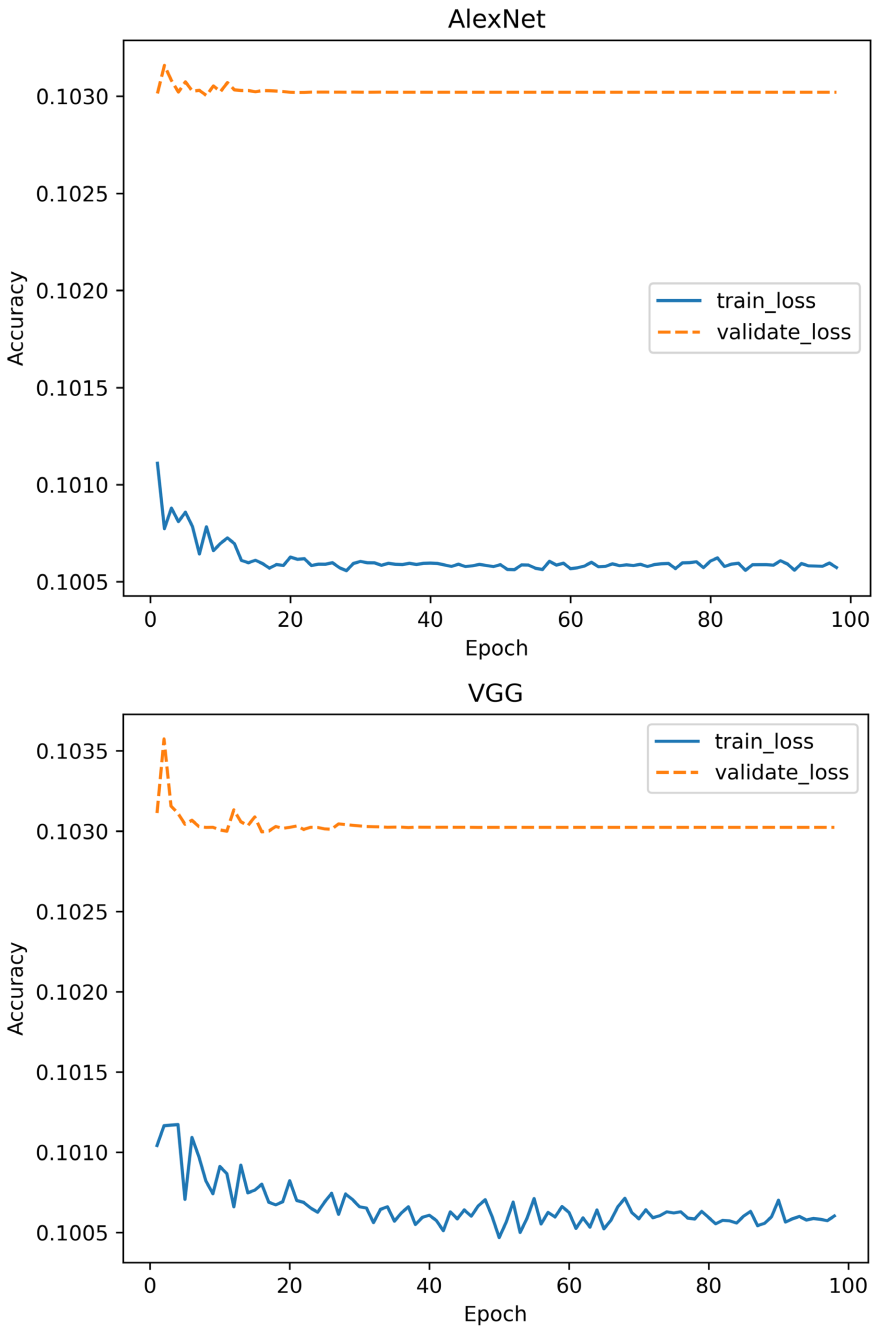

Figure 7. The confusion matrix of AlexNet and VGG for human recognition.

these two models cannot perform real activity recognition for FMCW signal. Therefore, DenseNet has the best recognition performance among these neural networks.

Although we obtained the expected identification result, there are still some issues that need to be addressed when using FMCW. Firstly, we should invite more participants and design more activities to evaluate the model. This will help us to validate the model's performance on a larger and more diverse dataset. Secondly, we should conduct more experiments in more environments to evaluate the model's performance. This will help us to determine how well the model can generalize to different environments and scenarios. To sum up, we should choose more neural network models to implement human activity recognition using a FMCW signal.

3. Conclusion

Currently, human-machine interaction has become a rapidly growing research area in artificial intelligence applications. FMCW-based human activity recognition has also increased attention due to its device-free pattern and good identification accuracy. This paper studies human activity recognition of five activities using DenseNet and FMCW signals. Specifically, we utilized the DenseNet model to implement human activity identification based on the collected point cloud. After data preprocessing, we fed the point cloud data into the neural network model to extract features and perform activity recognition. We achieved human activity recognition of five activities and obtained 100% test performance. Differently, the AlexNet and VggNet models almost did not converge after twenty epochs. The result shows that FMCW point cloud can effectively leverage DenseNet to extract features and implement activity classification. In addition, the results also validate that the FMCW signal has its characteristics compared with other wireless signals though we can consider these signals as a picture. These conclusions will help us to further improve the accuracy and robustness of the recognition system and make it more practical for real-world applications.

Acknowledgments

The work is funded by the foundation of the Innovation and Entrepreneurship Training Program for College Students (X202210424040).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Liu, J., Liu, H., Chen, Y., Wang, Y. and Wang, C. (2017) Wireless Sensing for Human Activity: A Survey. *IEEE Communications Surveys & Tutorials*, **22**, 1629-1645. https://doi.org/10.1109/COMST.2019.2934489

[2] Fu, B., Damer, N., Kirchbuchner, F. and Kuijper, A. (2020) Sensing Technology for Human Activity Recognition: A Comprehensive Survey. *IEEE Access*, **8**, 83791-83820. https://doi.org/10.1109/ACCESS.2020.2991891

[3] Gupta, N., Gupta, S.K., Pathak, R.K., Jain, V., Rashidi, P. and Suri, J.S. (2022) Human Activity Recognition in Artificial Intelligence Framework: A Narrative Review. *Artificial Intelligence Review*, **55**, 4755-4808. https://doi.org/10.1007/s10462-021-10116-x

[4] Thariq Ahmed, H.F., Ahmad, H. and Aravind, C.V. (2020) Device Free Human Gesture Recognition Using Wi-Fi CSI: A Survey. *Engineering Applications of Artificial Intelligence*, **87**, Article ID: 103281. https://doi.org/10.1016/j.engappai.2019.103281

[5] Huang, Y.C., Li, W., Dou, Z.Y., Zou, W.T., Zhang, A.Y. and Li, Z. (2022) A Activity Recognition Based on Millimeter-Wave Radar by Fusing Point Cloud and Range-Doppler Information. *Signals*, **3**, 266-283.https://doi.org/10.3390/signals3020017

[6] Halperin, D., Hu, W.J., Sheth, A. and Wetherall, D. (2011) Tool Release: Gathering 802.11n Traces with Channel State Information. *ACM SIGCOMM Computer Communication Review*, **41**, 53. https://doi.org/10.1145/1925861.1925870

[7] Wang, W., Liu, A.X., Shahzad, M., Ling, K. and Lu, S. (2015). Understanding and Modeling of WiFi Signal Based Human Activity Recognition. *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, Paris, 7-11 September 2015, 65-75. https://doi.org/10.1145/2789168.2790093

[8] Wang, Y., Liu, J., Chen, Y.Y., Gruteser, M., Yang, J. and Liu, H.B. (2014) E-Eyes: Device-Free Location-Oriented Activity Identification Using Fine-Grained WiFi Signatures. *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, Maui, 7-11 September 2014, 617-628. https://doi.org/10.1145/2639108.2639143

[9] Wang, Y., Wu, K. and Ni, L.M. (2017) WiFall: Device-Free Fall Detection by Wireless Networks. *IEEE Transactions on Mobile Computing*, **16**, 581-94. https://doi.org/10.1109/TMC.2016.2557792

[10] Zhu, D., Pang, N., Li, G. and Liu, S. (2016) WiseFi: Activity Localization and Recognition on Commodity Off-the-Shelf WiFi Devices. 2016 *IEEE 18th International Conference on High Performance Computing and Communications*; *IEEE 14th International Conference on Smart City*; *IEEE 2nd International Conference on Data Science and Systems* (*HPCC*/*SmartCity*/*DSS*), Sydney, 12-14 December 2016, 581-594. https://doi.org/10.1109/HPCC-SmartCity-DSS.2016.0085

[11] Li, X.Y., He, Y. and Jing, X.J. (2019) A Survey of Deep Learning-Based Human Activity Recognition in Radar. *Remote Sensing*, **11**, Article 1068 https://doi.org/10.3390/rs11091068

[12] Ding, C., Hong, H., Zou, Y., Chu, H., Zhu, X., Fioranelli, F., Le Kernec, J. and Li, C. (2019) Continuous Human Motion Recognition with a Dynamic Range-Doppler Trajectory Method Based on FMCW Radar. *IEEE Transactions on Geoscience and Remote Sensing*, **57**, 6821-6831. https://doi.org/10.1109/TGRS.2019.2908758

[13] Ding, C., Zhang, L., Chen, H., Hong, H., Zhu, X. and Fioranelli, F. (2023) Sparsity-Based Human Activity Recognition with Pointnet Using a Portable FMCW Radar. *IEEE Internet of Things Journal*, **10**, 10024-10037. https://doi.org/10.1109/JIOT.2023.3235808

[14] Saeed, U., Shah, S.Y., Shah, S.A., Ahmad, J., Alotaibi, A.A., Althobaiti, T., *et al.* (2021) Discrete Human Activity Recognition and Fall Detection by Combining FMCW RADAR Data of Heterogeneous Environments for Independent Assistive Living. *Electronics*, **10**, Article 2237. https://doi.org/10.3390/electronics10182237

[15] Wang, B., Guo, L., Zhang, H. and Guo, Y.-X. (2020) A Millimetre-Wave Radar-Based Fall Detection Method Using Line Kernel Convolutional Neural Network. *IEEE Sensors Journal*, **20**, 13364-13370. https://doi.org/10.1109/JSEN.2020.3006918

[16] Wang, B., Zheng, Z. and Guo, Y.-X. (2022) Millimeter-Wave Frequency Modulated Continuous Wave Radar-Based Soft Fall Detection Using Pattern Contour-Confined Doppler-Time Maps. *IEEE Sensors Journal*, **22**, 9824-9831. https://doi.org/10.1109/JSEN.2022.3165188

[17] Alizadeh, M., Shaker, G., Almeida, J.C.M.D., Morita, P.P. and Safavi-Naeini, S. (2019) Remote Monitoring of Human Vital Signs Using Mm-Wave FMCW Radar. *IEEE Access*, **7**, 54958-54968. https://doi.org/10.1109/ACCESS.2019.2912956

[18] Turppa, E., Kortelainen, J.M., Antropov, O. and Kiuru, T. (2020) Vital Sign Monitoring Using FMCW Radar in Various Sleeping Scenarios. *Sensors*, **20**, Article 6505. https://doi.org/10.3390/s20226505

[19] Sacco, G., Piuzzi, E., Pittella, E. and Pisa, S. (2020) An FMCW Radar for Localization and Vital Signs Measurement for Different Chest Orientations. *Sensors*, **20**, Article 3489. https://doi.org/10.3390/s20123489

[20] Wang, F., Zhang, F., Wu, C., Wang, B. and Liu, K.J.R. (2021) ViMo: Multiperson Vital Sign Monitoring Using Commodity Millimeter-Wave Radio. *IEEE Internet of Things Journal*, **8**, 1294-1307. https://doi.org/10.1109/JIOT.2020.3004046

[21] Krizhevsky, A., Sutskever, I. and Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*, **25**, 84-90. https://doi.org/10.1145/3065386

[22] Simonyan, K. and Zisserman, A. (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition. *The 3rd International Conference on Learning Representations* (*ICLR* 2015), San Diego, 7-9 May 2015, 1-14.

[23] Huang, G., Liu, Z., Pleiss, G., van der Maaten, L. and Weinberger, K.Q. (2022) Convolutional Networks with Dense Connectivity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **44**, 8704-8716. https://doi.org/10.1109/TPAMI.2019.2918284

[24] LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. *Nature*, **521**, 436-444. https://doi.org/10.1038/nature14539

[25] Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks*, **61**, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003

[26] Aggarwal, C.C. (2018) Neural Networks and Deep Learning. Springer, Cham. https://doi.org/10.1007/978-3-319-94463-0

[27] Dong, S., Wang, P. and Abbas, K. (2021) A Survey on Deep Learning and Its Applications. *Computer Science Review*, **40**, Article ID: 100379. https://doi.org/10.1016/j.cosrev.2021.100379

[28] Zhang, A., Lipton, Z., Li, M. and Smola, A. (2023) Dive into Deep Learning (PyTorch). Posts & Telecom Press, Beijing.