EAI Endorsed Transactions
on e-Learning
Review Article EAI.EU

Gesture Recognition Based on Deep Learning: A Review

Meng Wu[1,*]

[1] School of Physics and Information Engineering, Jiangsu Second Normal University, Nanjing, Jiangsu 210016, P R China

Abstract

Gesture recognition is an important and inevitable technology in modern times, its appearance and improvement greatly improve the convenience of people's lives, but also enrich people's lives. It has a wide range of applications in various fields. In daily life, it can carry out human-computer interaction and the use of smart home. In terms of medical treatment, it can help patients to recover and assist doctors to carry out experiments. In terms of entertainment, it allows users to interact with the game in an immersive manner. This paper chooses three technologies that deep learning plays a more prominent role in gesture recognition, namely CNNs, LSTM and transfer learning based on deep learning. They each have their own advantages and disadvantages. Because of the different principles of use, different techniques have different roles, such as CNNs can carry out feature extraction, LSTM can deal with long time series, transfer learning can transfer what is learned from another task to this task. Select different practical technologies according to different application scenarios and make improvements in real time in practical applications. Gesture recognition based on deep learning has the advantages of good accuracy, robustness, and real-time implementation, but it also bears the disadvantages of huge economic and time costs and high hardware requirements. Despite some challenges, researchers continue to optimize and improve the technology, and believe that in the future, gesture recognition technology will be more mature and valuable.

Keywords: Gesture Recognition; Deep Learning; CNNs; LSTM

Received on 23 February 2024, accepted on 02 March 2024, published on 07 March 2024

doi: 10.4108/eetel.5191

*Corresponding author. Email: jysw316@jssnu.edu.cn

1. Gesture Recognition

Gesture recognition is a process of converting human hand or body movements into understandable commands or interactive signals through computer vision and machine learning techniques [1, 2]. Gestures include human hand and body movements, and even facial movements. Gesture recognition technology can be applied in many fields, such as human-computer interaction, virtual reality, healthcare, and smart homes.

1.1 Human-computer interaction

Gesture recognition can be used as a natural and intuitive way of human-computer interaction [3-5], so that people can control computers or other devices through gestures [6, 7]. Gesture recognition can replace traditional input devices such as mouse [8] and keyboard, providing a more flexible and intuitive interactive experience, so that people can interact with computers more freely. For example, through gestures, you can control the operation of the computer, browse the web, play music [9] and so on.

1.2 Virtual reality

Gesture recognition can be used in the field of virtual reality with important application value [10-12], enabling users to operate the virtual environment or interact with virtual objects through gestures [13, 14]. For example, in virtual reality games [15, 16], users can control the actions of game characters through gestures, improving the sense of realism and interaction of the game, so that users are more immersed in the game.

1.3 Health care

Gesture recognition is also widely used in the field of health care, for rehabilitation training [17, 18], surgical assistance and patient monitoring [19]. For example, gesture recognition can be used in rehabilitation training to help rehabilitated patients with physical recovery therapy of movement. In addition, gesture recognition can also be used in surgical navigation and medical operations such as assisting doctors in surgery, which can improve the accuracy and safety of surgery.

1.4 Smart home

Gesture recognition can be used in smart home systems, allowing users to control home devices [20] and systems through gestures, such as adjusting lighting, temperature and volume, etc., improving the intelligence and convenience of furniture [21, 22].

In general, the purpose of gesture recognition is to achieve a natural, intuitive and intelligent way of computer interaction, providing better interactive experience and application value in different fields. With the continuous progress of computer hardware and algorithms, the real-time performance of gesture recognition has been significantly improved. Modern computer algorithms and hardware acceleration technology enable gesture recognition to be identified and tracked at high speed in real-time scenarios, meeting the needs of practical applications. It is believed that gesture recognition will also show its potential and application value in more fields. Some application fields and corresponding examples are collected in Table 1.

Table 1. Applications of gesture recognition

Application field	Example(s)
Human-computer interaction	Control the operation of the computer, browse the web, play music through gestures
Virtual reality	Control the actions of game characters through gestures
Health care	(i) Help rehabilitated patients with physical recovery therapy (ii) Assist doctors in surgery
Smart home	Control home devices and systems through gestures

This paper starts with the introduction of basic gesture recognition, and then expounds the deep learning techniques for gesture recognition, including CNNs, LSTM and transfer learning. Then this paper focuses on the general working steps of gesture recognition, from data collection to model optimization, but different problems require different solutions, and specific working steps should be selected according to the actual situation. Finally, this paper studies the advantages and disadvantages of gesture recognition based on deep learning and looks forward to its broad future development prospects. It is believed that with the continuous progress and improvement of technology, gesture recognition will be applied to more fields.

2. Deep Learning for Gesture Recognition

Gesture recognition based on deep learning has made remarkable progress in recent years. The emergence of deep learning models [23-25] has greatly improved the accuracy and robustness of gesture recognition. At present, gesture recognition based on deep learning mainly uses convolutional neural networks (CNNs) [26] and long short-term memory network (LSTM) for feature extraction and classification. These models can automatically learn the features of gesture images and action sequences, and train and optimize the model through training data.

In general, gesture recognition based on deep learning continues to evolve in algorithms and techniques, making significant progress. With the improvement of hardware devices and computing power, as well as the increasing demand for gesture recognition applications, deep learning gesture recognition technology is expected to play a greater role in practical applications.

2.2 Convolutional Neural Networks (CNNs)

CNNs are deep learning models that are particularly suitable for image processing and play an important role in gesture recognition [27-29]. CNNs can be used to extract features from gesture images. Through multi-layer convolution and pooling operation, CNNs can automatically learn the spatial structure and local features of gesture images, to realize gesture classification and recognition. CNNs have the following characteristics.

1. Feature extraction: CNNs can automatically extract features [30] from images through convolution layer and pooling layer, which are of great significance for gesture recognition [31]. These features can be the edge, texture, shape and other features of gesture images, which can effectively represent the key information of gestures. By learning these features, CNNs can transform gesture images into high-dimensional feature vectors, to realize effective representation and recognition of gestures, and provide a basis for subsequent classification and recognition tasks [32].
2. Hierarchical representation: CNNs can gradually build hierarchical representation of images through multi-layer convolution and pooling operations. This hierarchical representation can help CNNs capture features of different levels in the image, from low-level edges and textures to high-level shapes and structures, to better represent the structure and content of gesture images and improve the accuracy of gesture recognition [33-35].
3. Weight sharing: The convolutional layer in CNNs uses a weight sharing mechanism [36], which means that the same set of weights are used when the convolutional kernel slides across the entire image [37]. This kind of weight sharing can greatly reduce the number of model parameters and improve the training efficiency of the model. For gesture recognition tasks, weight sharing can better capture local features in the image and is very useful for local shape and texture extraction in gesture recognition.
4. Translation invariance: CNNs have certain translation invariance [38], which means that even if the gesture is translated in the image, CNNs can still recognize the same gesture. This nature makes CNNs more robust in handling gesture recognition tasks and able to cope with changes in the position of gestures in the image. The characteristics and advantages of CNNs are summarized below in Figure 1.

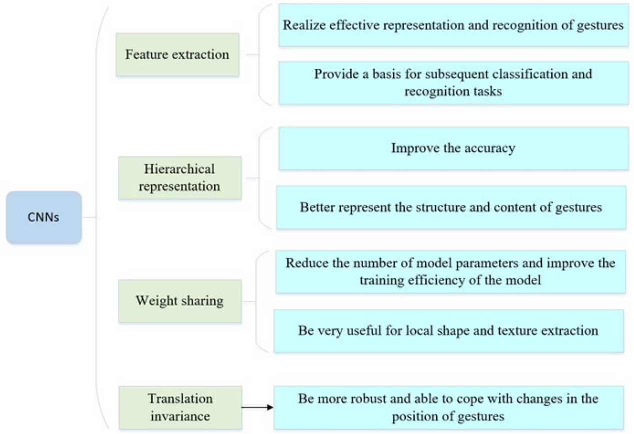

Figure 1. The characteristics and advantages of CNNs

To sum up, CNNs play an important role in gesture recognition, providing powerful modeling capabilities and recognition performance for gesture recognition tasks through features extraction, hierarchical representation, weight sharing, translation invariance and other features. Gesture recognition based on CNNs is a very promising technology, which can be widely used in many fields.

2.2 Long Short-Term Memory Network (LSTM)

Recurrent neural network (RNN) is a deep learning model suitable for sequential data processing [39]. In gesture recognition, RNN can be used to process gesture sequence data, such as the time series of gestures. Through the structure of the loop, RNN can capture the timing information in the gesture sequence, to realize the continuous recognition and tracking of gestures. LSTM is a special type of RNN, which can deal with long sequence data effectively [40] and can solve the problem of gradient disappearance and gradient explosion in traditional RNN. In gesture recognition, LSTM can be used to process long gesture sequences and improve the modeling ability of time sequence information [41].

By introducing a gating mechanism to control the flow of information, LSTM can effectively process long sequence data [42]. It contains a memory cell [43] that can store and read information and control the flow of information through a gate unit [44]. Gating units include input gates, forget gates, and output gates, which learn parameters to decide whether to allow information to flow in, out, or remain unchanged. The composition of LSTM is shown in Figure 2.

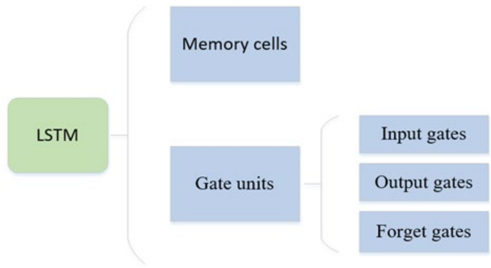

Figure 2. The composition of LSTM

The main functions in gesture recognition are as follows:

1. Process time series data: Gesture recognition involves time series data, that is, the changes of gestures at different points in time. LSTM can process time sequence data effectively [45] and can capture the change rule of gesture in time dimension through its internal memory unit and gating mechanism. This enables LSTM to better model and understand the timing features of gestures in gesture recognition tasks.
2. Long-term dependence on modeling: In gesture recognition, some gesture features may be observed over a longer time. Traditional RNN are prone to gradient disappearance or gradient explosion when faced with long-term dependence problems. However, LSTM can effectively deal with the long-term dependency problem through its gating mechanism, especially the forgetting gate and the input gate, so that the model can better capture the long-term dependency in the gesture sequence [46].
3. Model context information: Gestures in gesture recognition are often interrelated, and the action of the previous gesture may affect the action of the subsequent gesture. Through its memory unit and gating mechanism, LSTM can effectively model the context information and pass the information of the previous gesture to the subsequent gesture. This enables LSTM to better understand the context in gesture sequences and improve the accuracy of gesture recognition. Some advantages and their functions are shown in Table 2.

Table 2. The description of LSTM

Advantage	Function(s)
Process time series data	Better model and understand the timing features
Long-term dependence on modeling	Better capture the long-term dependency
Model context information	(i) Better understand the context in gesture sequences (ii) Improve the accuracy

In gesture recognition, LSTM can receive gesture sequence as input and establish the mapping relationship between gesture and corresponding action through learning. LSTM can effectively capture the timing information in gesture sequences, and store important information through memory units for processing longer gesture sequences and improve the modeling ability of the model for timing information. In conclusion, LSTM plays an important role in gesture recognition.

2.3 Transfer Learning for Deep Learning Models

Transfer learning for deep learning models [47] is the technique of transferring knowledge learned on one task to another related task [48, 49]. In gesture recognition, transfer learning can improve the accuracy and generalization of the model by pre-training the model on a large-scale data set and then fine-tuning it on a smaller one [50]. Specifically, the first several layers of the existing image recognition model can be used as feature extractors, and the gesture recognition data set can be input into these layers to obtain the corresponding feature representation. Some additional layers can then be added for training on gesture recognition tasks. Since the existing model has learned some general feature representations in other tasks, these feature representations are also helpful for gesture recognition tasks, which can speed up the convergence rate and improve the accuracy of the model [51]. The working steps of transfer learning is shown in Figure 3.

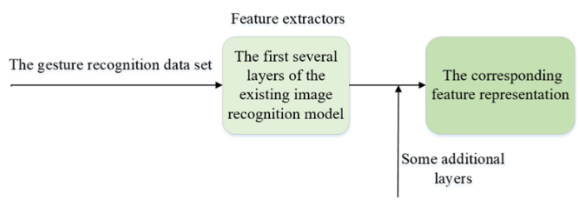

Figure 3. The working steps of transfer learning

Transfer learning can also solve the problem of small gesture recognition datasets. Due to the high acquisition cost of gesture recognition datasets, there is often only a limited amount of data available for training. In this case, the use of transfer learning can leverage large-scale data sets of existing tasks to extract common features, so that training can still be effective in the case of small samples [52, 53].

In short, transfer learning can speed up model training and improve accuracy in gesture recognition, especially in the case of small data sets. So that gesture recognition models can better adapt to the characteristics of different gesture data sets.

These deep learning techniques play an important role in gesture recognition, improving the accuracy and robustness of gesture recognition [54]. With the continuous development of deep learning technology, we can also expect more innovative deep learning models to be applied to gesture recognition.

3. Gesture Recognition Steps via Deep Learning

Deep learning-based gesture recognition is a technology that uses deep learning algorithms to recognize human gestures, which can automatically recognize and understand the role and intent of human gestures by capturing ethical images or video data through cameras or other sensors [55]. Gesture recognition based on deep learning typically involves the following steps:

1. Data collection: Use cameras [56] or other sensors to collect gesture data sets [57, 58], including images or videos of different gestures. The dataset should contain enough samples to represent a variety of gesture gestures and different environmental conditions.
2. Data preprocessing: The collected gesture data is pre-processed to improve the performance of the model. Common pre-processing steps include image or video scaling, cropping, grayscale, normalization, etc.
3. Build a deep learning model: Select an appropriate deep learning model architecture, such as CNNs or RNN, and design and adjust the model according to the characteristics of the data set. Using the model to extract the features of gesture data [59], for example, CNNs can extract the texture and shape features of the image.
4. Model training: Use marked gesture data sets to train deep learning models. In the training process, the weight and parameters of the model are adjusted by backpropagation algorithm and optimization algorithm, so that the model can accurately recognize gestures. After the result of the gesture is obtained, the result is identified and classified.
5. Model optimization: According to the training results, the model is optimized, such as adjusting the hyperparameters of the model, adding training data, using data enhancement technology, etc., to further improve the performance of the model [60, 61]. These steps are shown in Figure 4.

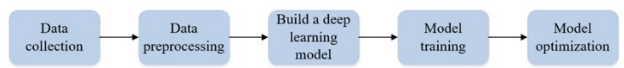

Figure 4. Steps of gesture recognition

It should be noted that the above steps are only the general process of gesture recognition, and the specific implementation methods and technical details may vary depending on project requirements and data set characteristics.

4. The Advantages of Gesture Recognition via Deep Learning

Compared with traditional gesture recognition methods, gesture recognition based on deep learning has the following advantages:

a) High accuracy: Deep learning models can be trained with a large amount of training data to improve the accuracy of gesture recognition. Deep learning models can learn more complex features and patterns to better distinguish between different gestures.

b) Robustness: Deep learning models have strong robustness to factors such as light changes and background interference. Deep learning models can learn abstract feature representations to reduce their dependence on concrete environments.

c) Scalability: Deep learning models can improve gesture recognition performance by adding training data and adjusting network structure. At the same time, deep learning models can be combined with other deep learning techniques, such as object detection, pose estimation, etc., to improve the ability of gesture recognition. With the introduction of more data and more complex network structures, the performance of gesture recognition can continue to improve [62, 63].

d) Real-time [64]: The deep learning model can recognize gestures in a short period of time through high-performance hardware acceleration and optimization algorithms to realize real-time gesture recognition applications. This is very important for some application scenarios that require fast response, such as virtual reality, smart homes, and so on.

e) End-to-end learning [65]: Traditional gesture recognition methods often require manual design of feature extractors and classifiers. Gesture recognition based on deep learning can learn features and classifiers directly from the original image data through end-to-end training, avoiding the tedious process of manually designing features.

f) Portability: Deep learning models can be applied to different gesture recognition tasks through transfer learning. By training a good model on one task, it can be fine-tuned on another related task, thus speeding up model training and improving model performance. Table 3 shows the collected advantages of gesture recognition.

Table 3. The description of advantages of gesture recognition

Advantage	Specific description
High accuracy	Train a large amount of data
Robustness	Reduce dependence on concrete environments.
Scalability	Add training data and adjust network structure
Real-time	Use high-performance hardware acceleration and optimize algorithms
End-to-end learning	Learn features and classifiers directly from the original image data
Portability	Use transfer learning

5.The Disadvantages of Gesture Recognition via Deep Learning

Although gesture recognition based on deep learning has advantages that are not comparable to other methods, it also has some disadvantages:

a) Large and demanding data: Deep learning models require a large amount of labeled data for training, especially for complex gesture recognition tasks, requiring more data to be able to accurately recognize gestures. Collecting and labeling large amounts of gesture data is a time-consuming and expensive task. Meanwhile, deep learning models have high requirements for the quality and accuracy of input data. If the input data is noisy or distorted, the recognition performance may be degraded.

b) Long training time: The training of deep learning models usually requires a lot of computing resources and time. Training an accurate gesture recognition model can take hours or even days, which increases the time cost of developing and deploying a gesture recognition system.

c) High hardware requirements: The training and inference of deep learning models usually require high-performance hardware, such as GPU. For real-time gesture recognition applications, high-performance hardware devices are required to meet the computing requirements, which limits the ability to perform real-time gesture recognition on low-end devices.

d) Difficult to explain: Deep learning models are often considered black box models [66], and it is difficult to explain how the model makes recognition decisions. This makes it difficult to understand errors in the model and improve recognition accuracy. Disadvantages of gesture recognition are shown in Figure 5.

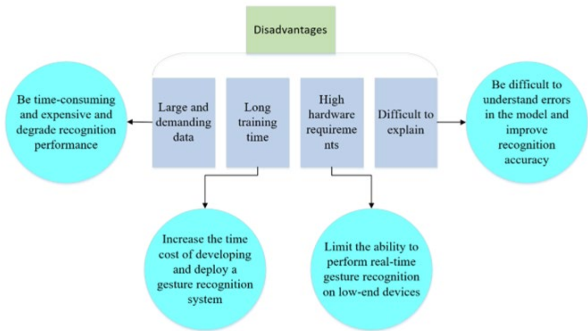

Figure 5. The disadvantages of gesture recognition

To address these issues, the researchers have proposed approaches such as data enhancement, transfer learning, and semi-supervised learning to reduce the need for labeled data. In addition, gesture recognition based on deep learning faces some challenges [67]. For example, the diversity and complexity of gestures make the model need to have strong generalization ability. The real-time requirement makes the model need to recognize and track efficiently in a short time. For the recognition of continuous action, the model needs to have a certain time series modeling ability. To address these challenges, the researchers propose many improved approaches. For example, the introduction of attention mechanisms to focus on important areas in gesture images, combine sensor data such as inertial measurement unit and image data for multi-modal gesture recognition. Then, reinforcement learning method is used to optimize the performance of gesture recognition.

6. Conclusion

Gesture recognition based on deep learning has made remarkable progress in recent years. Gesture recognition is a technology that identifies human intentions by analyzing and understanding human actions represented by gestures. It has broad prospects and is widely used in many fields such as virtual reality and human-computer interaction. There are many deep learning techniques, in the field of gesture recognition, the typical ones are CNNs and LSTM and transfer learning. As one of the most important technologies in modern times, CNNs have broad application prospects in various fields, including gesture recognition. It uses convolution layer, pooling layer and full connection layer to extract and classify image features and completes the task of gesture recognition well. Gesture recognition requires the use of cameras or sensors to collect data, so attention to these external devices is also indispensable. Each technology has its own advantages, but also has its own shortcomings, we believe that with the continuous progress and development of technology, deep learning technology in gesture recognition application will be more, the performance will be better.

Acknowledgment

We thank all the anonymous reviewers for their hard reviewing work.

Funding

The research work was supported by the open project of State Key Laboratory of Millimeter Waves (Grant No. K202218).

References

[1] K. R. Pyun, K. Kwon, M. J. Yoo, K. K. Kim, D. Gong, W. H. Yeo, et al., "Machine-learned wearable sensors for real-time hand-motion recognition: toward practical applications," National science review, vol. 11, pp. nwad298-nwad298, 2024.
[2] N. Fadel and E. I. A. Kareem, "Detecting Hand Gestures Using Machine Learning Techniques," Ingénierie des Systèmes d'Information, vol. 27, 2022.
[3] D. Fu, "Human-computer Interaction Gesture Control Recognition for High-performance Embedded AI Computing Platform," Advances in Computer and Communication, vol. 4, 2023.
[4] T. Min, "Gesture Detection and Recognition Based on Pyramid Frequency Feature Fusion Module and Multiscale Attention in Human-Computer Interaction," Mathematical Problems in Engineering, vol. 2021, 2021.
[5] L. Zhihan, P. Fabio, D. Qi, L. Jaime, and S. Houbing, "Deep Learning for Intelligent Human–Computer Interaction," Applied Sciences, vol. 12, pp. 11457-11457, 2022.
[6] J. Lesong, Z. Xiaozhou, and X. Chengqi, "Non-trajectory-based gesture recognition in human-computer interaction based on hand skeleton data," Multimedia Tools and Applications, vol. 81, pp. 20509-20539, 2022.
[7] A. Mohammad, A. Abdullah, A. Gharbi, A. T. S, A. Saud, A. Dhahi, et al., "Human-computer Interaction Gesture Control Recognition for High-performance Embedded AI Computing Platform," Advances in Computer and Communication, vol. 4, 2023.
[8] N. Prasanth, K. Shrivastava, A. Sharma, A. Basu, R. A. Sinha, and S. P. Raja, "Gesture-based mouse control system based on MPU6050 and Kalman filter technique," International Journal of Intelligent Systems Technologies and Applications, vol. 21, pp. 56-71, 2023.
[9] C. Alexandre, M. Luciano, R. Miriam, and B. Gilberto, "Musical Control Gestures in Mobile Handheld Devices: Design Guidelines Informed by Daily User Experience," Multimodal Technologies and Interaction, vol. 5, pp. 32-32, 2021.
[10] Y. Chen, Q. Li, C. Huang, C. Ye, Y. Li, and R. Lu, "Research on the Virtuality-Reality Interaction Based on Dynamic Gesture Recognition in Augmented Reality System," Advanced Science Letters, vol. 7 2012, pp. 468-472(5), 2012.
[11] F. Zhao, W. Jinlong, and N. Taile, "Research and Application of Multifeature Gesture Recognition in Human-Computer Interaction Based on Virtual Reality Technology," WIRELESS COMMUNICATIONS & MOBILE COMPUTING, vol. 2021, 2021.
[12] L. Jiang, X. Yu, and L. Wang, "A brief analysis of gesture recognition in VR," SID Symposium Digest of Technical Papers, vol. 51, pp. 190-195, 2020.
[13] H. Asim, M. Sebastian, and P. Andrew, "How good are virtual hands? Influences of input modality on motor tasks in virtual reality," Journal of Environmental Psychology, vol. 92, 2023.
[14] P. Anastasia, G. Geoffrey, B. Domna, and W. Sebastian, "Using virtual reality to assess gesture performance deficits in schizophrenia patients," Frontiers in psychiatry, vol. 14, pp. 1191601-1191601, 2023.
[15] W. Peng, W. Yue, B. Mark, Y. Huizhen, X. Peng, and L. Yanhong, "BeHere: a VR/SAR remote collaboration system based on virtual replicas sharing gesture and avatar in a procedural task," Virtual reality, vol. 27, pp. 21-22, 2023.
[16] B. Ganguly, P. Vishwakarma, Rahul, and S. Biswas, "Gesture based Virtual Reality Implementation for Car Racing Game employing an Improved Neural Network," International Journal of Engineering Research and, vol. V8, 2019.
[17] L. Wen, W. Shunxin, L. Simou, Z. Xiyang, Z. Xiaobo, Q. Hao, et al., "Gesture Recognition System Using Reduced Graphene Oxide-Enhanced Hydrogel Strain Sensors for Rehabilitation Training," ACS applied materials & interfaces, vol. 15, 2023.
[18] G. Kai, O. Mostafa, L. Jingxin, S. A. Maged, Y. Hongbo, and E. Mahmoud, "Empowering Hand Rehabilitation with AI-Powered Gesture Recognition: A Study of an sEMG-Based System," Bioengineering (Basel, Switzerland), vol. 10, 2023.
[19] D. M. Doriana, D. S. Elisa, and V. Giovanni, "Embodying Language through Gestures: Residuals of Motor Memories Modulate Motor Cortex Excitability during

Abstract Words Comprehension," Sensors, vol. 22, pp. 7734-7734, 2022.
[20] P.Vogiatzidakis and P.Koutsabasis, "Mid-air gesture control of multiple home devices in spatial augmented reality prototype," Multimodal Technologies and Interaction, vol. 4, pp. 1-22, 2020.
[21] Z. Sheng, S. Zening, Z. Wenjie, S. Xu, X. Linghui, J. Bo, et al., "An EMG-based wearable multifunctional Eye-control glass to control home appliances and communicate by voluntary blinks," Biomedical Signal Processing and Control, vol. 86, 2023.
[22] J. Dai, "Gesture Recognition Based Smart Home Control System," Journal of Electronic Research and Application, vol. 3, pp. 7-8, 2019.
[23] Y.-D. Zhang, "Advances in Multimodality Data Fusion in Neuroimaging," Information Fusion, vol. 76, pp. 87-88, 2021.
[24] S. Wang, "Advances in data preprocessing for biomedical data fusion: an overview of the methods, challenges, and prospects," Information Fusion, vol. 76, pp. 376-421, 2021.
[25] J. Wang, "A Review of Deep Learning on Medical Image Analysis," Mobile Networks & Applications, vol. 26, pp. 351-380, 2021.
[26] S. Arooj, S. Altaf, S. Ahmad, H. Mahmoud, and A. S. N. Mohamed, "Enhancing sign language recognition using CNN and SIFT: A case study on Pakistan sign language," Journal of King Saud University - Computer and Information Sciences, vol. 36, pp. 101934-, 2024.
[27] A. H. Victoria and G.Maragatham, "Gesture recognition of radar micro doppler signatures using separable convolutional neural networks," Materials Today: Proceedings, vol. 80, pp. 1961-1964, 2023.
[28] B. H. Lee, T. Kim, and D. H. Oh, "3D Virtual Reality Game with Deep Learning-based Hand Gesture Recognition," Journal of the Korea Computer Graphics Society, vol. 24, pp. 41-48, 2018.
[29] C. C. d. Santos, J. L. A. Samatelo, and R. F. Vassallo, "Dynamic gesture recognition by using CNNs and star RGB: A temporal information condensation," Neurocomputing, vol. 400, pp. 238-254, 2020.
[30] A. A. Barbhuiya, R. K. Karsh, and R. Jain, "CNN based feature extraction and classification for sign language," Multimedia Tools and Applications, vol. 80, pp. 1-19, 2020.
[31] J. P. Sahoo, S. P. Sahoo, S. Ari, and S. K. Patra, "RBI-2RCNN: Residual Block Intensity Feature using a Two-stage Residual Convolutional Neural Network for Static Hand Gesture Recognition," Signal, Image and Video Processing, vol. 16, pp. 1-9, 2022.
[32] A. O. Tolulope, A. Adewale, K. Faiq, and H. S. Rafay, "FM-ModComp: Feature Map Modification and Hardware–Software Co-Comparison for secure hardware accelerator-based CNN inference," Microprocessors and Microsystems, vol. 100, 2023.
[33] S.-H. Wang and M. A. Khan, "VISPNN: VGG-inspired stochastic pooling neural network," Computers, Materials & Continua, vol. 70, pp. 3081-3097, 2022.
[34] S.-H. Wang, K. M. Attique, and G. Vishnuvarthanan, "Deep rank-based average pooling network for COVID-19 recognition," Computers, Materials, & Continua, vol. 70, pp. 2797-2813, 2022.
[35] X. Guo, "A Survey on Machine Learning in COVID-19 Diagnosis," Computer Modeling in Engineering & Sciences, vol. 130, pp. 23-71, 2022.
[36] S. Ziyi, L. Handong, Q. Jinwu, Z. Zhen, and Z. Lunwei, "Hand Gesture Recognition Based on sEMG Signal and Convolutional Neural Network," International Journal of Pattern Recognition and Artificial Intelligence, vol. 35, 2021.
[37] N. Cui, "Applying Gradient Descent in Convolutional Neural Networks," Journal of Physics: Conference Series, vol. 1004, pp. 012027-012027, 2018.
[38] O.S.Kayhan and J. C. v. Gemert, "On translation invariance in CNNs: Convolutional layers can exploit absolute spatial location," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 14262-14273, 2020.
[39] A. B. Ibrahimm, A. Hira, M. N. Al, A. Abdulwahab, A. Abdullah, A. S. S, et al., "Smart Home Automation-Based Hand Gesture Recognition Using Feature Fusion and Recurrent Neural Network," Sensors (Basel, Switzerland), vol. 23, 2023.
[40] T. Alejandro, J. Juan, C. T. Juan, P. Alejandro, L. Alexandro, and C. R. Alexandre, "LSTM Recurrent Neural Network for Hand Gesture Recognition Using EMG Signals," Applied Sciences, vol. 12, pp. 9700-9700, 2022.
[41] R. Zahra, M. Malik, and D. Frédéric, "Autonomous gesture recognition using multi-layer LSTM networks and laban movement analysis," International Journal of Knowledge-based and Intelligent Engineering Systems, vol. 26, pp. 289-297, 2023.
[42] M. U. Rehman, F. Ahmed, M. A. Khan, U. Tariq, F. A. Alfouzan, N. M. Alzahrani, et al., "Dynamic Hand Gesture Recognition Using 3D-CNN and LSTM Networks," Computers, Materials & Continua, vol. 70, pp. 4675-4690, 2022.
[43] H. Giannis, R. Gerasimos, and A. Ioannis, "A Modified Long Short-Term Memory Cell," International journal of neural systems, vol. 33, pp. 2350039-2350039, 2023.
[44] A. N. Moustafa and W. Gomaa, "Gate and common pathway detection in crowd scenes and anomaly detection using motion units and LSTM predictive models," Multimedia Tools and Applications, vol. 79, pp. 1-40, 2020.
[45] W. Weina, S. Jiapeng, and J. Huxidan, "Fuzzy inference-based LSTM for long-term time series prediction," Scientific Reports, vol. 13, pp. 20359-20359, 2023.
[46] F. Liu, L. Zhang, and Z. Jin, "Modeling programs hierarchically with stack-augmented LSTM," The Journal of Systems & Software, vol. 164, pp. 110547-110547, 2020.
[47] Z. Zhen, L. Shilong, W. Yanyu, S. Wei, and Z. Yuhui, "Online cross session electromyographic hand gesture recognition using deep learning and transfer learning," Engineering Applications of Artificial Intelligence, vol. 127, 2024.
[48] G. Alfonso, M. Delfina, Z. Rocco, C. Carmine, and L. Nicola, "Touchscreen gestures as images. A transfer learning approach for soft biometric traits recognition," Expert Systems With Applications, vol. 219, 2023.
[49] K. Peiqi, L. Jinxuan, F. Bingfei, J. Shuo, and B. S. Peter, "Wrist-worn Hand Gesture Recognition while Walking via Transfer Learning," IEEE journal of biomedical and health informatics, vol. PP, 2021.
[50] Z. Shenyilang, F. Yinfeng, W. Jiacheng, J. Guozhang, and L. Gongfa, "Transfer Learning Enhanced Cross-Subject

Hand Gesture Recognition with sEMG," Journal of Medical and Biological Engineering, vol. 43, pp. 672-688, 2023.
[51] R. Yu and G. Guoying, "Deep transfer learning-based adaptive gesture recognition of a soft e-skin patch with reduced training data and time," Sensors and Actuators: A. Physical, vol. 363, 2023.
[52] Q. Bu, G. Yang, X. Ming, T. Zhang, J. Feng, and J. Zhang, "Deep transfer learning for gesture recognition with WiFi signals," Personal and Ubiquitous Computing, vol. 26, pp. 1-12, 2020.
[53] W. Fei, H. Ronglin, and J. Ying, "Research on Transfer Learning of Vision-based Gesture Recognition," International Journal of Automation and Computing, vol. 18, pp. 1-10, 2021.
[54] A. R.Patil and S.Subbaraman, "Performance analysis of static hand gesture recognition approaches using artificial neural network, support vector machine and two stream based transfer learning approach," International Journal of Information Technology, vol. 14, pp. 1-12, 2021.
[55] K. Yang, M. Kim, Y. Jung, and S. Lee, "Hand Gesture Recognition Using FSK Radar Sensors," Sensors, vol. 24, 2024.
[56] Z. Yinger, W. Zhouyi, L. Peiying, P. Yang, W. Yuting, Z. Liufang, et al., "Hand gestures recognition in videos taken with a lensless camera," Optics express, vol. 30, pp. 39520-39533, 2022.
[57] R. Dmitry, I. Denis, and R. Elena, "Audio-Visual Speech and Gesture Recognition by Sensors of Mobile Devices," Sensors, vol. 23, pp. 2284-2284, 2023.
[58] H. Pingao, W. Hui, W. Yuan, G. Yanjuan, Y. Wenlong, G. Chao, et al., "In-Situ Measuring sEMG and Muscle Shape Change with a Flexible and Stretchable Hybrid Sensor for Hand Gesture Recognition," IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society, vol. PP, 2022.
[59] E. N. Rubén and E. B. Marco, "Analysis and Evaluation of Feature Selection and Feature Extraction Methods," International Journal of Computational Intelligence Systems, vol. 16, 2023.
[60] H. Ashraf, A. Waris, S. O. Gilani, U. Shafiq, J. Iqbal, E. N. Kamavuako, et al., "Optimizing the performance of convolutional neural network for enhanced gesture recognition using sEMG," Scientific reports, vol. 14, pp. 2020-2020, 2024.
[61] W. S. Josiah, T. Shiva, W. Richard, M. Yiorgos, and T. Murat, "Improved Static Hand Gesture Classification on Deep Convolutional Neural Networks Using Novel Sterile Training Technique," IEEE ACCESS, vol. 9, pp. 10893-10902, 2021.
[62] Y. D. Zhang and S. C. Satapathy, "Fruit category classification by fractional Fourier entropy with rotation angle vector grid and stacked sparse autoencoder," Expert Systems, vol. 39, 2022.
[63] L. Yang, "EDNC: ensemble deep neural network for Covid-19 recognition," Tomography, vol. 8, pp. 869-890, 2022.
[64] S.Rajko, G.Qian, T.Ingalls, and J.James, "Real-time gesture recognition with minimal training requirements and on-line learning," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2007.
[65] R. Georg, S. Moritz, F. Tim, and B. Luca, "7 [formula omitted]J/inference end-to-end gesture recognition from dynamic vision sensor data using ternarized hybrid convolutional neural networks," Future Generation Computer Systems, vol. 149, pp. 717-731, 2023.
[66] B. Anton, H. Andreas, O. Emilia, K. Kimmo, and P. Kai, "Explaining any black box model using real data
," Frontiers in Computer Science, vol. 5, 2023.
[67] H. Zhou, D. Wang, Y. Yu, and Z. Zhang, "Research Progress of Human–Computer Interaction Technology Based on Gesture Recognition," Electronics, vol. 12, 2023.