Gesture Recognition of sEMG Signals based on Deep Learning Framework

CHENG YANG, CHENXUAN ZHANG*
School of Artificial Intelligence, Liangjiang,
Chongqing University of Technology,
Chongqing, 401135,
P.R.CHINA
Corresponding Author

Abstract: - Surface electromyography (sEMG) signal has been a hot research topic in the field of human-computer interaction technology in recent years, It is not disturbed by environmental factors such as light, temperature, and humidity, and has the advantages of high precision, fast response and non-intrusiveness. Through the application of sEMG signals, the intelligent device can accurately judge the person's movement intention. Convolutional neural networks (CNNs) and long short-term memory networks (LSTM) are considered to have better performance on sequence data. In this paper, three deep learning frameworks (1-dimensional CNN, 2-dimensional CNN, and CNN-LSTM) are used for the gesture recognition task of continuous sEMG signals and evaluated for recognition performance separately. The results show that the 2D-CNN has the best recognition effect, which achieved average recognition accuracy of 90.36%. The average recognition accuracy of the CNN-LSTM and 1D-CNN is 89.37% and 80.21%, respectively. In addition, the time-domain sliding window segmentation method was used to process the EMG signal sequences to ensure the objectivity of the evaluation processes of CNN-LSTM.

Key-Words: - Deep learning, sEMG signals, Gesture recognition, Time domain characteristic parameters of the sliding window, Human-computer interaction, Multi-class classification.

Received: March 22, 2024. Revised: August 17, 2024. Accepted: November 3, 2024. Published: December 5, 2024.

1 Introduction

Surface electromyography is a bioelectric signal of muscle movement recorded by surface electrodes, which can reflect the functional status of nerves and muscles. With the development of information technology and artificial intelligence, the use of computer processing of surface EMG signals will help to better interpret the relationship between signals and physiological performance or to predict and identify various actions, which has an important role in the field of medical rehabilitation and sports health. There are already many researchers working on the application of computer and artificial intelligence technologies to the processing and recognition of sEMG data. In 2014, [1] used a PSD featuring four gesture actions as the input support vector machine for recognition, and the recognition rate reached 91.97%. [2], used RMS, HIST, and other features to classify and recognize the gestures provided by Ninapro, and the highest accuracy of 52 types of actions was 75.27% by using a random forest neural network. In 2016, [3] used ANN to identify 72 eigenvalues extracted from sEMG signals, and the recognition accuracy of 16 hand movements could reach 87.8%. In 2017, [4] used RMS and other features to classify and recognize 52 gestures of 27 people in the Ninapro database, with the highest classification accuracy of 69.13%. [5], applied the deep learning model to gesture recognition, using the method based on the deep convolutional network, through the comparison experiment of gesture actions provided by the Ninapro database, and the trained classifier classified the input EMG images frame by frame. In 2018, [6] used a moving average energy method to segment the data of 8-channel sEMG signals from subjects, extracted the mean absolute value (MAV) of time-domain features, and then used the dynamic time warping (DTW) algorithm to classify four commonly used gestures. [7], combined the root mean square (RMS) and integrated EMG values in the time-domain features extracted from the sEMG signal and used PNNs to classify the four commonly used gestures. In 2023, [8], achieved an accuracy of 91.27% on the ninapro's 9-class classification task by using large convolutional kernels and large pooling in CNNs to reduce the feature map size, and by using a cascaded network to mitigate feature loss. [9] proposed an association model based on a two-level structure of CNN with an accuracy of 85.5% for 52 gesture classifications on Ninapro.

This study aims to realize and evaluate deep learning frameworks for gesture recognition based on sEMG signals, explore three current main architectures, namely 1-dimensional CNN, 2-dimensional CNN, and hybrid network of CNN and LSTM, and determine the model that is best suited for sEMG signal data.

2 Material and Methods

Please, leave two blank lines between successive sections as here.

Mathematical Equations must be numbered as follows: (1), (2), …, (99) and not (1.1), (1.2),…, (2.1), (2.2),… depending on your various Sections.

2.1 Dataset Description and Processing

The surface EMG signal data used in this paper were obtained from the Ninapro database, which is intended to be a benchmark database in the field of non-invasive adaptive hand prosthetics, [10]. The dataset 2 and 3 were selected for study, which includes complete subjects and amputees in total 40, respectively. Each subject was asked to perform 12 basic gestures, and the surface electromyography signal generated by each gesture was recorded. The 12 basic gestures and rest movements are shown in Figure 1.

Fig. 1: 12 basic gestures and rest diagrams

Since the raw data does not distinguish between the signals of each gesture, the data needs to be cleaned before the model can be applied. Clustering algorithms can be used to distinguish the characteristics of different gesture signals, such as amplitude, number, etc. Rectifying via RMS makes the input data correspond to a time window of 150ms, spanning all available electrode measurements, i.e., 10 electrodes for delays. This choice is very much in line with the work that is typically done, and only the time window that reflects an effective gesture will be analyzed. The original signal and the clustered gesture signal are shown in Figure 2 (Appendix). To facilitate the evaluation of the model in the later stage, the dataset must be divided, and Table 1 shows the distribution of datasets used for 1D-CNN and 2D-CNN training and testing.

Table 1. The dataset distribution of 1D-CNN and 2D-CNN is divided into 70%, 15%, and 15% respectively for the training set, validation set, and test set

Data set	Percentage of data(%)	1D-CNN data volume	2D-CNN data volume
Training	70.0	43268	40580
Validation	15.0	9272	8696
Test	15.0	9272	8696

2.2 Three Pre-Selected Model Structures

Based on previous literature results, three different network architectures were designed: 1-dimensional convolutional neural network (1D-CNN), 2-dimensional convolutional neural network (2D-CNN), and network model mixed with convolutional network and long and short memory network (CNN-LSTM). In 1D-CNN, the input dimension was 10, which corresponds to 10 gestures representing numbers, the convolutional kernel size was set to 1x3, a total of 4 convolutional layers were used, and the activation function has used Tanh. The 2D-CNN is the same as 1D-CNN, but the kernel size was set to 3x3, and added BatchNorm layer and Dropout layer. For CNN-LSTM, we replaced the fully connected layer with the LSTM structure based on 2D-CNN, and other parameters were the same as those of 2D-CNN. The structure diagrams of the three networks are shown in Appendix in Figure 3, Figure 4 and Figure 5.

2.3 Experimental Environment

Hardware devices used in this study are AMD Ryzen 7 5800H with 32GB of memory, NVIDIA GeForce RTX 3050 graphics card with 4G video memory, and Windows 11-x64 operating system. Visual Studio Code is used as the experimental platform, Python language is used for programming, and Pytorch (2.0.1+cu118) is used as the basic framework for deep learning. All data were loaded onto the GPU via CUDA for model training.

3 Results

3.1 The Training Performance of Each Model

The hyperparameters used in the training of models were the same, with a learning rate of 0.1, a momentum of 0.8, and a batch number of 256. The recognition performance changes of the models during the training process were recorded, which were represented by Cross-entropy loss and accuracy. Table 2 shows the average recognition accuracy of models on the test set. In Appendix, Figure 6, Figure 7 and Figure 8 illustrate the performance of three methods during training, respectively. It is worth noting that although the recognition accuracy of CNN-LSTM is not the highest, it performs best in the early training stage, and the convergence speed of the model is relatively smoother and faster.

Table 2. The gesture recognition results on the test dataset

Methods	Average recognition accuracy(%)
1D-CNN	80.21
2D-CNN	90.36
CNN-LSTM	89.37

3.2 A Time-Domain Sliding Window Segmentation Method for CNN-LSTM Evaluation

For the generation of training samples of two-dimensional data, the time-domain sliding window segmentation method was adopted, and the recognition accuracy of the generated samples under different sampling repetition rates (overlap) was studied. Due to the limitation of experimental research hardware, we only carried out experimental research on CNN-LSTM with the recognition effect in the middle. Table 3 shows the recognition accuracy results.

With the decrease of the repetition rate, the recognition accuracy of CNN-LSTM on the test set shows a downward trend, which indicates the importance of step size based on the time domain sliding window to a certain extent, which reflects the resolution/response rate of the gesture recognition instrument in practical application, and the higher the response rate of the instrument, the lower the corresponding value step, and the higher the recognition accuracy.

Table 3. The recognition accuracy of CNN-LSTM with different coverage rates is based on the time-domain sliding window segmentation method

Model	Repetition rate(%)	The average recognition accuracy of the train(%)	Average recognition accuracy of tests(%)
CNN-LSTM	75.0	90.13	87.41
	50.0	91.29	81.87
	25.0	92.71	77.35
	0.0	90.93	76.65

4 Conclusion

In this study, three classical deep learning frameworks were used to recognize gesture movements using surface EMG signal data. The results show that 2D-CNN shows higher accuracy when processing time-scale data such as EMG signals, and the recognition accuracy of gestures reaches 90.36%. However, CNN-LSTM has better performance in convergence speed, and the average accuracy also reaches 89.37%, which means that LSTM is better at processing time series while maintaining a high level of recognition. In addition, by applying the time-domain sliding window segmentation method, it can be observed that the performance of CNN-LSTM is affected by the sampling repetition rate, which has an enlightening effect on data acquisition engineering, because the response rate of the sampling device may improve the performance of the software.

Nowadays, surface EMG signal gesture recognition is a research field that has attracted much attention in the field of human-computer interaction, and the continuous development of deep learning technology has also greatly promoted research in this field. With the continuous progress and optimization of deep learning technology, we expect to achieve more in-depth research results in the field of surface EMG signal gesture recognition in the future.

References:

[1] Kim, J., Cho, D., Lee, K. J., & Lee, B. (2015). A real-time pinch-to-zoom motion detection by means of a surface EMG-based human-computer interface. *Sensors*, *15*(1), 394-407. https://doi.org/10.3390/s150100394

[2] Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Hager, A. M., Elsig, S., Giatsidis, G., Bassetto, F., & Müller, H. (2014). Electromyography data for non-invasive

naturally-controlled robotic hand prostheses. *Scientific Data*, *1*, Article 140053. https://doi.org/10.1038/sdata.2014.53

[3] Luh, G.-C., Ma, Y.-H., Yen, C.-J., & Lin, H.-A. (2016). Muscle-gesture robot hand control based on sEMG signals with wavelet transform features and neural network classifier. In *2016 International Conference on Machine Learning and Cybernetics (ICMLC)* (pp. 627-632). Jeju, Korea(South). https://doi.org/10.1109/ICMLC.2016.7872960

[4] Pizzolato, S., Tagliapietra, L., Cognolato, M., Reggiani, M., Müller, H., & Atzori, M. (2017). Comparison of six electromyography acquisition setups on hand movement classification tasks. *PLOS ONE*, *12*(10), e0186132. https://doi.org/10.1371/journal.pone.0186132

[5] Du, Y., Jin, W., Wei, W., Hu, Y., & Geng, W. (2017). Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation. *Sensors*, *17*(3), 458. https://doi.org/10.3390/s17030458

[6] Xiaoyu, X., & Chee-Jie, L. (2018). Gesture recognition method for EMG signals based on DTW algorithm. *Computer Engineering and Applications*, *54*(5), 132-137. https://doi.org/10.3778/j.issn.1002-8331.1610-0061

[7] Qingli, W., Wei, X., Weiqiang, L., Zhenchao, S., & Li, Z. (2018). Gesture recognition based on PNN. *Sensors and Microsystems*, *37*(8), 16-18. https://doi.org/10.13873/j.1000-9787(2018)08-0016-03

[8] Qureshi, M. F., Mushtaq, Z., Rehman, M. Z. U., & Kamavuako, E. N. (2023). E2CNN: An efficient concatenated CNN for classification of surface EMG extracted from the upper limb. *IEEE Sensors Journal*, *23*(8), 8989-8996. https://doi.org/10.1109/JSEN.2023.3255408

[9] Fan, J., Wen, J., & Lai, Z. (2023). Myoelectric pattern recognition using Gramian angular field and convolutional neural networks for muscle-computer interface. *Sensors (Basel)*, *23*(5), 2715. https://doi.org/10.3390/s23052715

[10] Connolly, C. (2008). Prosthetic hands from Touch Bionics. *Industrial Robot*, *35*(4), 290-293. https://doi.org/10.1108/01439910810876364

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)
Cheng Yang contributed all the experiments and the first draft, and Chenxuan Zhang has completed the final draft.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself
No funding was received for conducting this study.

Conflict of Interest
The authors have no conflicts of interest to declare.

APPENDIX

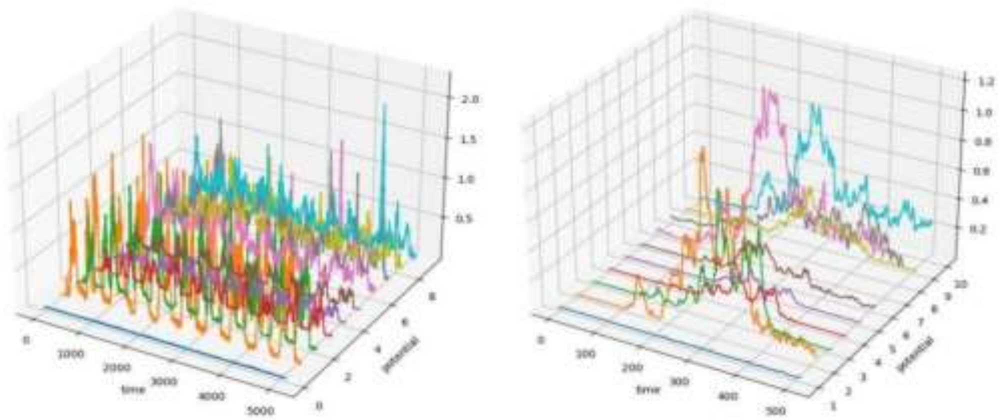

Fig. 2: Diagram with resting raw signal (left) and single gesture raw signal (right)

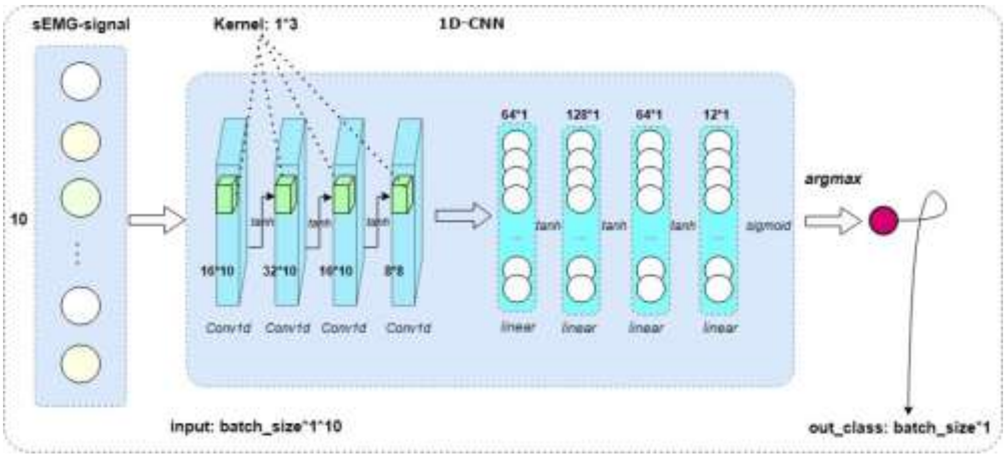

Fig. 3: Schematic diagram of the structure of 1D-CNN

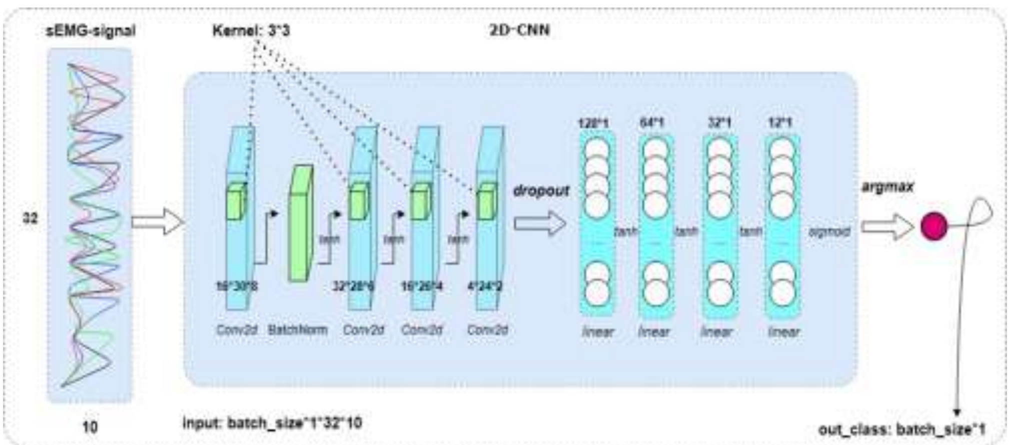

Fig. 4: Schematic diagram of the structure of 2D-CNN

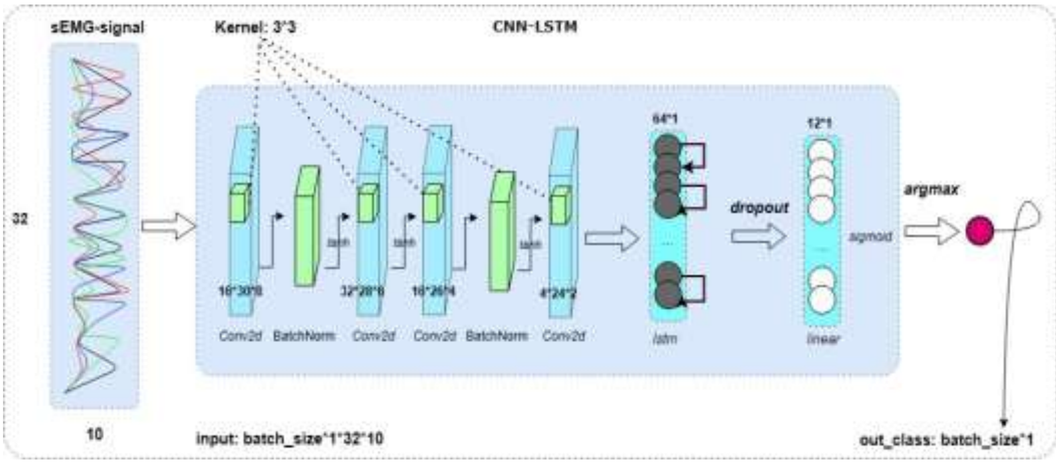

Fig. 5: Schematic diagram of the structure of CNN-LSTM

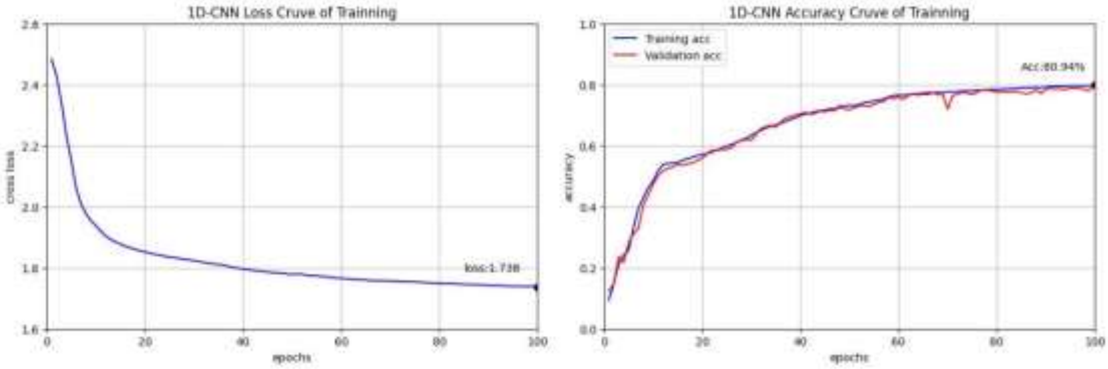

Fig. 6: Changes in the training loss (left) and validation accuracy (right) of the 1D-CNN model

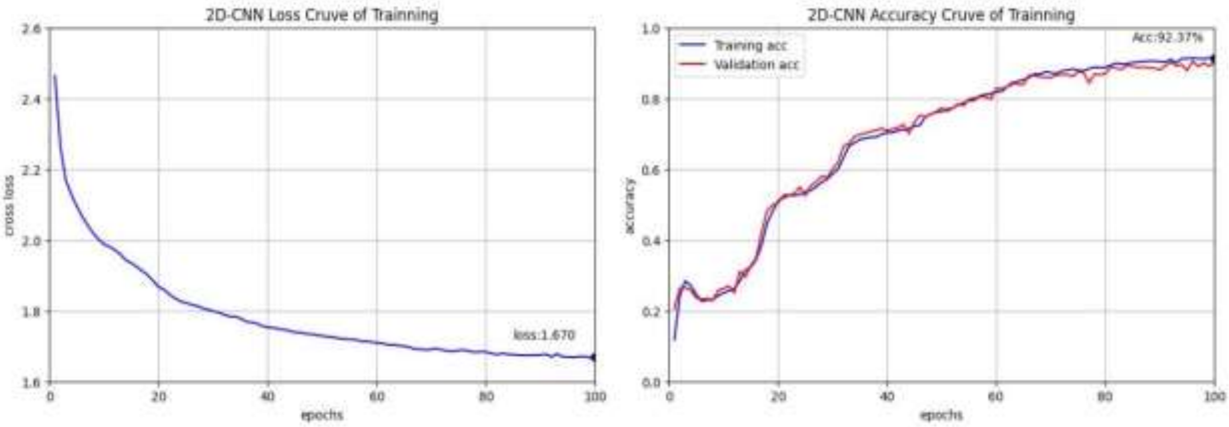

Fig. 7: Changes in the training loss (left) and validation accuracy (right) of the 2D-CNN model

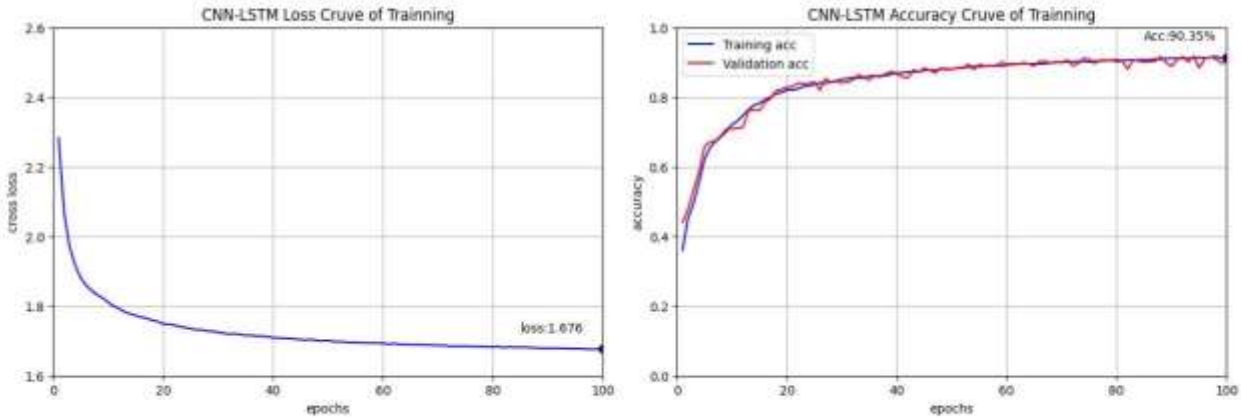

Fig. 8: Changes in the training loss (left) and validation accuracy (right) of the CNN-LSTM model