2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
October 11-14, 2020. Toronto, Canada

Feature Fusion of sEMG and Ultrasound Signals in Hand Gesture Recognition

Jia Zeng[1], Yu Zhou[1], Yicheng Yang[1], Jiaole Wang[2], and Honghai Liu[13*]

[1]State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai, China
[2]State Key Laboratory of Robotics and System, School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen, China
[3]School of Computing, University of Portsmouth, Portsmouth, UK
[*]honghai.liu@icloud.com

Abstract—**Multi-modal sensory fusion is believed to obtain higher accuracy in gesture recognition. Its difficulty lies in mining discriminative features and fusing features from different modalities. Surface electromyography(sEMG) and ultrasound signals are typical signal modalities in gesture recognition. It is expected that the fusion of them can take advantage of the complementarity of electrophysiological information and muscle morphology information. This paper proposed two kinds of feature fusion method. The one is concatenating the manual designed sEMG and ultrasound features, and the other is a convolutional neural network (CNN) based feature exaction and fusion method for sEMG and ultrasound signals. Eight able-bodied subjects were involved to participate in the experiments. In the experiments, four channels of sEMG and A-mode ultrasound signals corresponding to 20 gestures were collected synchronously to evaluate the proposed method. The experimental results demonstrated that the fusion sEMG-ultrasound feature always outperformed the separate sEMG or ultrasound feature regardless of the feature extraction method, and as for fusion sEMG-ultrasound feature, the CNN based method achieve a high accuracy** $(97.38 \pm 1.49\%)$ **in 20 gestures, which surpassed the method of concatenating the manual designed features and applying machine learning algorithm (LDA, KNN, SVM).**

I. Introduction

In recent years, neural prosthesis has been leveraged to improve the daily lives of amputees. In the application scenario of prosthetic hand control, gesture recognition based on biosignals [1] plays a key role. Due to non-invasive and easy to acquirement, surface electromyography (sEMG) signal becomes the most commonly used human-machine interface in gesture recognition. In recent years, research on prosthetic limb control based on sEMG pattern recognition has been widely carried out [2]. Current research shows that with appropriate feature extraction and pattern recognition algorithms, for 10 types of wrist and gesture actions, an offline recognition accuracy of more than 95% can be obtained [3]. However, when the gestures to be recognized has a greater number of categories and is more sophisticated, the recognition accuracy of sEMG will be faced with severe challenges. Based on 12-channel sEMG data of 50 gestures from 40 subjects provided by NinaPro DB2 database [4], Zhai et al. [5] achieved a classification accuracy of 70% on KNN algorithm and 78.71% on SVM algorithm. The recognition rate drops a lot. Deep learning algorithm is widely used as it automatically extracts effective features through neural networks, and many researchers managed to introduce it to improve the performance of sEMG signals in the classification of sophisticated gestures. On NinaPro DB2 dataset, Ding et al. [6] proposed a CNN based classification model C-B1PB2 to attain a higher classification accuracy of 78.86%, and Hu et al. [7] designed a attention-based hybrid CNN-RNN architecture to reach 82.2%. These latest studies proved that the recognition accuracies obtained by deep neural networks were better than the methods based on manual feature extraction and machine learning.

Although sEMG has made a lot of achievements in the field of pattern recognition for prosthesis control, its defects such as crosstalk and low spatial resolution limit its ability in decoding accurately dexterous finger motions [8]. As ultrasound has the advantages of high detection accuracy and high spatial and temporal resolution [9], [10], it can detect the morphological changes of the surface and deep muscles, which make it better to reflect the muscle contractions and map them to gestures. Ultrasound is divided into A-mode ultrasound and B-mode ultrasound. Although B-mode could attain excellent gesture recognition accuracy [11], [12], the employed B-mode ultrasound equipment is too bulky to be wearable. A-mode ultrasound, which has been designed to be portable [13], could be an ideal human-machine interface. Yang et al. [14] apply LDA algorithm on eight channels of A-mode ultrasound data and achieve average classification accuracy of 98.83% across 11 finger motions. In another study, Xia et al. [15] applied SVM algorithm on four channels of A-mode ultrasound data and attained 84.22% across 20 complicated gestures while the same method on four channels of sEMG data achieved only 68.59%. Ultrasound signal performs much better than sEMG signal in sophisticated gesture recognition accuracy.

To further improve the performance of human machine interface, many researches focused on multi-modal signal fusion. Guo et al. [16] designed hybrid sEMG/NIRS sensor system. The recognition rate using sEMG/NIRS fusion feature set $(> 97\%)$ will be significantly improved compared to separate

This work was supported by the National Natural Science Foundation of China (No. 51575338, 51575407).

sEMG or NIRS feature set ($< 90\%$). [17] developed a fusion sEMG/mMMG system, and the average classification accuracy of mMMG, EMG and combined EMG-mMMG are 81.6%, 90.8% and 95.1%, which demonstrates the potential complementarity between sEMG and mMMG. [18] showed that the combination of sEMG and force sensitive resistors (FSR) could reach the accuracy of 96.05% for 21 subtle gestures. The above research verified that the recognition accuracy of the human-machine interface benefits from multi-modal signal fusion. The fusion of sEMG and ultrasound is expected to take advantage of the complementarity of electrophysiological information and muscle morphology information and finally further improve the recognition accuracy.

The main contribution of this study is proposing two effective fusion method for sEMG-ultrasound signal and demonstrating the feasibility of sEMG-ultrasound feature fusion.

II. Dataset Collection and Evaluation Metric

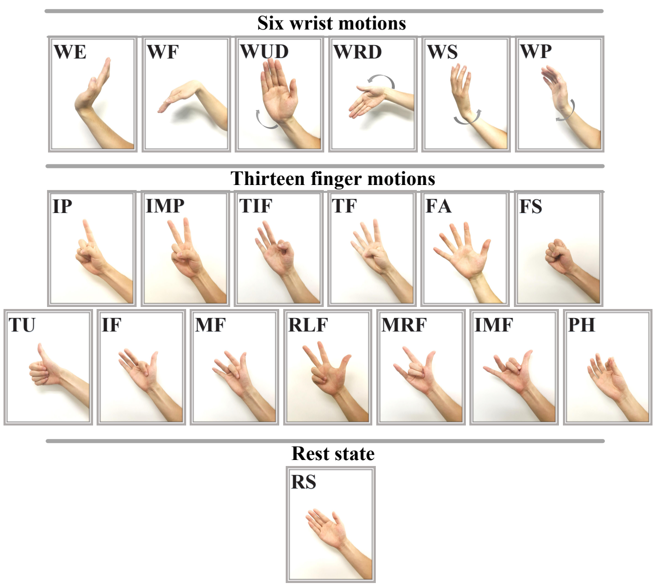

Fig. 1. The gestures paradigm of experiments in this paper. There are 20 types of gestures, including 6 wrist movements, 13 finger movements and a rest state (RS). Six wrist movement includes: wrist extension (WE), wrist flexion (WF), wrist ulnar deviation (WUD), wrist radial deviation (WRD), wrist supination (WS) and wrist pronation (WP); Thirteen finger movement includes: index point (IP), index and middle finger point (IMP), thumb and index finger flexion (TIF), thumb flexion (TF), fingers abduction (FA), fist (FS), thumb up (TU), index flexion (IF), middle finger flexion (MF), ring and little finger flexion (RLF), middle and ring finger flexion (MRF), index and middle finger flexion (IMF) and pinch (PH).

To evaluate the performance of sEMG and ultrasound signal in gesture recognition, eight able-bodied subjects were recruited to participate in the experiment. The subjects were asked to perform 20 gestures in order as shown in Fig.1 according to screen guidance, with four channels of sEMG signal and A-mode ultrasound(AUS) signal recorded synchronously. The four channels of sEMG signal and AUS signal were collected by a four-channel hybrid sEMG/AUS sensor system, relevant details of which can be found in [15]. The sEMG electrode and the ultrasound probe shared the same acquisition position and the muscles involved in the collection point were flexor digitorum superficialis (FDS), flexor carpi ulnaris (FCU), extensor carpi ulnaris (ECU) and extensor digitorum (ED).

The involved twenty gestures are composed of 6 wrist motions, 13 finger motions and a rest state motion. The twenty consecutive actions are considered a trial. Each action lasts 5 seconds and there is no rest time between the two actions. The subjects were asked to repeat 8 trials, with an interval of 25s between each trial. Thus, 1000s of EMG and ultrasound data were collected from each subject.

The raw sEMG data was collected at a sampling rate of 1000 Hz, and the ultrasound data was are acquired at 10 frames per second. A series of sEMG data in a sliding window is regarded as a sample, and each frame of the ultrasound signal is regarded as a sample. In order to make the number of sEMG samples used for classification consistent with the number of ultrasound samples, a slide window with window size of 256ms and step of 100ms was adopted to divide sEMG samples. In addition, the data of first one and last one second were discarded because the action was in transition. And the 25s of data between each trial were discarded too. Thus, the number of sEMG samples and ultrasound samples was both 4800.

Since this paper does not involve inter-subject issue in bio-signal based gesture recognition which usually involves individual differences, the experimental dataset of these eight subjects were independently tested. In this article, 4-folder cross-validation was adopted to evaluate the predictive performance of the model. The data of each subject was divided into four groups. Take turns to select three of them as training set, and the remaining one group for test. The recognition accuracy on the testing set is defined as the ratio of the number of correctly classified samples to the total number of test samples, and the average recognition accuracy through cross-validation was employed as the classification metric of each subject. The final evaluation metric to classification model was the average recognition accuracy across all subjects.

III. Methods

This paper proposed two pipelines for sEMG and ultrasound signal processing and recognition for classification. The one is using manual designed features of sEMG and ultrasound according to the lone-term experience in the fields and applying machine learning algorithm such as Linear Discriminant Analysis (LDA), k-Nearest Neighbor (KNN) and Support Vector Machine (SVM). The other is leveraging convolutional neural network (CNN) to automatically extract sEMG and ultrasound features and project the extracted features to gestures.

A. *Manual Designed Features Method*

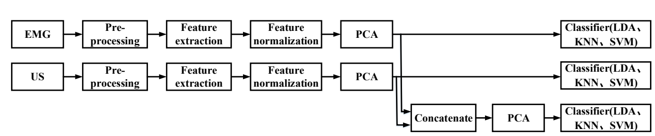

Fig. 2. The pipeline of extracting the manual designed features of sEMG and ultrasound respectively and combining the multi-modal features.

The pipeline of extracting manual designed features from sEMG/ultrasound signals and applying machine learning algorithms is shown in Fig.2. The process goes through preprocessing, feature extraction, feature normalization, principal component analysis (PCA) and classifier in turn. In view of circumstances in fusion sEMG and ultrasound signal, it is needed to concatenate features of sEMG and ultrasound after PCA and perform another dimensionality reduction by PCA.

1) Signal Preprocessing and Feature Extraction: The raw sEMG signal was comb-filtered with a frequency doubling of 50Hz and band-pass filtering with upper and lower cut-off frequencies of 20Hz and 500Hz. The statistical values of sEMG signals within the sliding window are calculated as the sEMG feature. TD-AR6 [19] feature was chosen as the sEMG signal feature in this paper as it is extensively used. The TD-AR6 feature includes four time domain feature (mean absolute value (MAV), waveform length (WL), zero crossing number (ZC), and slop sign changes (SSC)) and 6th order AR coefficient of autoregressive model (AR6). The TD-AR6 features across four channels were stacked finally, so the dimension of sEMG features is 40.

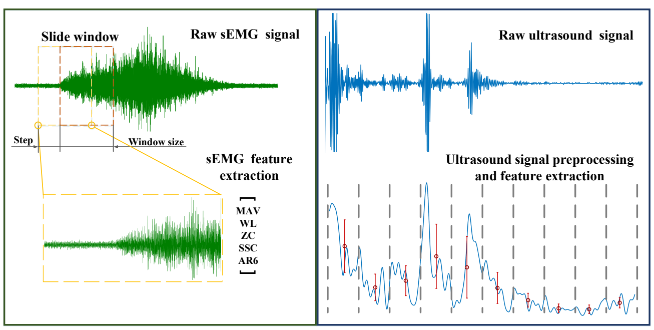

Fig. 3. The Illustration of signal preprocessing and feature extraction procedure for sEMG and ultrasound signal. The red error bars in the figure at the bottom right represent MSD features of ultrasound signal.

The raw ultrasound signal is preprocessed in order to remove noise and extract useful information. The preprocessing of ultrasound consists of time gain compensation, band-pass filtering, envelope detection, and log compression [14]. The first 20 dots and last 20 dots of each channel were discarded because they were unable to provide valid information and may have negative effects. For each channel, the remaining 960 dots were divided into 48 segmentations. The MSD feature [14] of ultrasound signal was extracted. The MSD feature represents the mean value and standard deviation calculated for each segmentation. The MSD features of all 48 segmentations across four channels were computed and stacked finally. Thus, the dimension of ultrasound feature is 384.

The original dimension of sEMG features is 40, and that of ultrasound features is 384. This paper employed PCA to reduce feature dimensions. PCA was firstly adopted in separate sEMG features and ultrasound features. And after feature combination, PCA was also adopted in conbined sEMG-ultrasound features.

2) Machine Learning Classifier: After extracting and optimizing the features, selecting an appropriate classifier is the next key step. The classifiers commonly used in the field of discrete gesture recognition are LDA, KNN, and SVM.

The involved LDA classifier is based on Gaussian assumptions and Bayesian decision criteria, and is currently one of the most widely used classifiers in the field of discrete gesture recognition based on physiological signals [20]. It should be noticed that LDA algorithm has no hyperparameter.

K-Nearest Neighbor algorithm is usually applied in classification problem. As for a test point, according to the selected distance formula, the nearest K points are filtered from the training data set and the test point will be classified as the class with the largest proportion of these K points. In this paper, the adopted metric that measures the distance between two sample points is Euclidean distance, as shown in Eq.1.

$$dist(\widetilde{X}, X) = \sqrt{\sum_{i=1}^{n} (\widetilde{X}_i - X_i)^2 / n} \tag{1}$$

where n represents the dimensions of feature. It should be noticed that the number of nearest points (K) selected is an important hyperparameter of the KNN algorithm.

Support vector machine (SVM) is a machine learning algorithm that seeks the maximum classification interval and the kernel function of SVM makes it a non-linear classifier. The discriminant of SVM model is defined as Eq.2

$$g(x) = w^T \cdot \psi(x) + b \tag{2}$$

where w^T represents weight matrix, $\psi(\cdot)$ represents kernel function, and b represents bias. The task of SVM algorithm is to find the optimal (w^T, b). In addition, to eliminate the impact of abnormal points on model training, slack variable ξ is introduced. The optimization goal translates to Eq.3

$$\begin{aligned} &\min_{w,b,\xi} \frac{1}{2}||\omega||^2 + C\sum_{i=1}^{m} \xi_i \\ &s.t. \quad \begin{cases} y_i(w^T \cdot \psi(x) + b) \geq 1 - \xi_i, \\ \xi_i \geq 0, i = 1, 2, ..., N. \end{cases} \end{aligned} \tag{3}$$

where N stands for the number of training samples, and penalty factor C was set to avoid ξ_i being too large. Penalty factor C is an important hyperparameter of the SVM algorithm. In this paper, the kernel function $\psi(\cdot)$ was chosen as Radial Basis Function (RBF) kernel.

3) Hyperparameter Tuning Strategy: The performance comparison of sEMG, ultrasound, and fusion sEMG-ultrasound mainly depends on the classification accuracy. In order to make the comparison fair, hyperparameter optimization should be carried out in PCA and the classifer.

When the classifier was chosen as KNN or SVM, the number of principal components, the hyperparameter of KNN classifier and SVM classifier were needed to be optimized jointly, which can be regarded as a multiple-layer optimization problem. During the hyperparameter tuning process, the search space for simultaneous PCA and classifier hyperparameter tuing is too large, and its time cost is unbearable. Particularly, LDA algorithm has no hyperparameter so it can be leveraged to tune the number of principal components . The tuning strategy in this paper is choosing LDA as classifier firstly, based on this tuning the number of principal components and fixing the number of principal components in the following tuning process, and finally tuing the hyperparameter of KNN classifier and SVM classifier with fixed number of principal components.

B. CNN Extracted Features Method

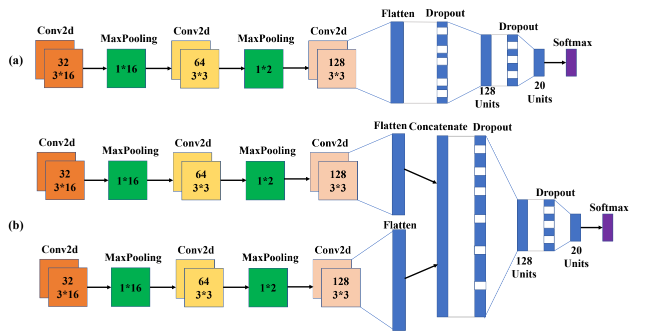

Fig. 4. The proposed CNN architecture.(a)The shared CNN architecture for separate sEMG or ultrasound signal to feature extraction and classification; (b)The CNN based feature extraction and feature fusion architecture for sEMG and ultrasound signals.

The adopted CNN architecture can automatically extract signal features and project them into gesture categories. The network structure as represented in Fig.4(a) is suitable for both sEMG and ultrasound signals. In order to satisfy the input requirements of the neural network, the sEMG and ultrasound signal were processed into images with size of 4x256 and 4x960 respectively, and normalized to [-1,1]. The network is composed of three convolutional layers, two max pooling layers, two fully-connected layers, two Dropout layers and Softmax output layer. The first convolutional layers have 32 2D filters of 3×16 with the stride of 1, followed by a MaxPooling layer of 1×16. The second convolutional layers have 64 2D filters of 3×3 with the stride of 1, followed by a MaxPooling layer of 1×2. The third convolutional layers have 128 2D filters of 3×3 with a valid padding. After that, the feature maps will be flatten into a feature vector, and after passing through two fully-connected layers and two Dropout layers, the probability of belonging to each gesture category is output through Softmax layer.

The proposed CNN based sEMG and ultrasound feature fusion architecture (EU-Net) is demonstrated in Fig.4(b). The proposed EU-Net model has two branches. The input of the first branch is an sEMG image, and the first branch extracts sEMG features. Correspondingly, the input of the second branch is an ultrasound image, and the function of the second branch is to extract ultrasound features. After extracting features by convolutional layers, the feature maps of sEMG and ultrasound are flatten into feature vectors and concatenated together. The subsequent process is consistent with that introduced in Fig.4(a) above.

IV. Experimental Results

A. Manual Designed Features Method

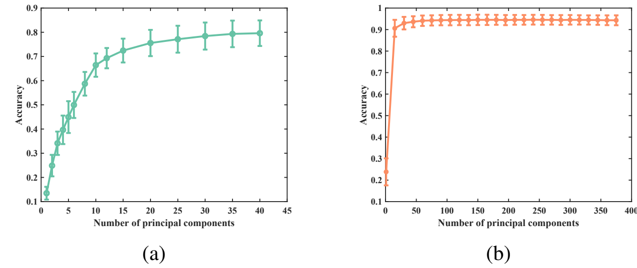

Fig. 5. Demonstration of PCA parameter tuning results: (a)The relationship between the classification accuracy of sEMG signals and the number of principal components; (b)The relationship between the classification accuracy of ultrasound signals and the number of principal components

1) Hyperparameter Tuning: The relationship between the classification accuracy of sEMG signals or ultrasound signals and the number of principal components are present in Fig.5(a) and Fig.5(b). It is verified by experiment that, to obtain the best classification accuracy, the dimension of EMG features after PCA should be consistent with the original dimension; the dimension of ultrasound features after PCA should be 240.

Combining the sEMG and ultrasound features after dimensionality reduction to form a fusion sEMG-ultrasound feature, and then performing PCA on the fusion sEMG-ultrasound feature. The relationship between the classification accuracy of sEMG-ultrasound signals and the number of principal components is shown as Fig.6. It could be found in this figure that the optimal PCA dimension was 270.

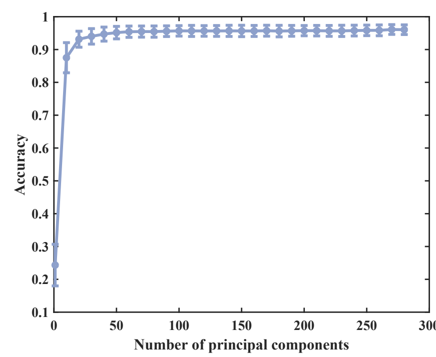

Fig. 6. The relationship between the classification accuracy of sEMG-ultrasound feature and the number of principal components.

After confirming the dimension reduction by PCA on the sEMG features, the ultrasound features and the fusion sEMG-ultrasound features, the confirmed number of principal components will be fixed all the time. After that the hyperparameters of the classifier are needed to be adjusted.

As involved classifer in this paper, LDA classifier has no hyperparameters, KNN classifier is mainly needed to adjust K value, and SVM classifier is mainly needed to adjust penalty factor C. The KNN hyperparameter tuning experimental results on separate sEMG signals, individual ultrasound signals and fusion of sEMG and ultrasound signals are shown in the Fig.7(a). The best K value can be selected as 35 for sEMG, 35 for ultrasound and 29 for fusion sEMG-ultrasound. The

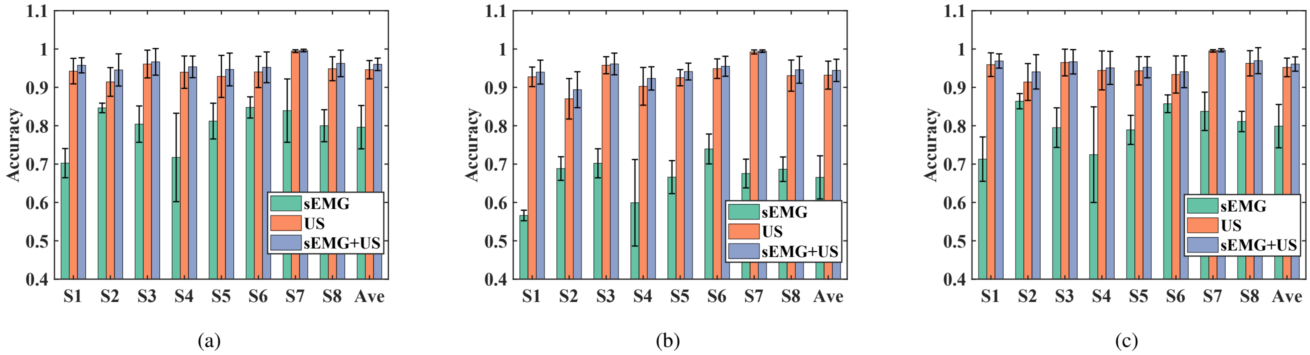

Fig. 8. The classification accuracy across the involved three datasets (individual sEMG, individual ultrasound(US), fusion sEMG and US):(a)LDA classifier; (b)KNN classifier;(c)SVM classifier.

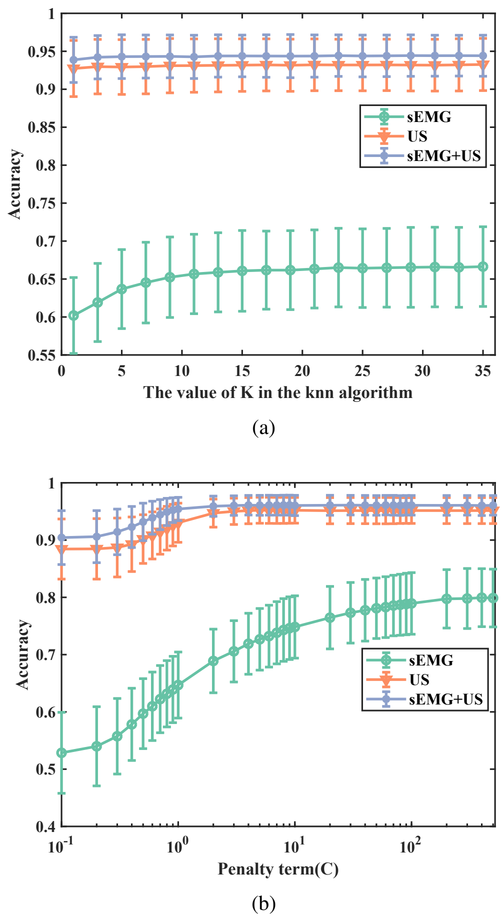

Fig. 7. The relationship between the classification accuracy of individual sEMG signal, individual ultrasound (US) signal or fusion of sEMG and ultrasound signals and the hyperparameter in the algorithm: (a) K value in KNN algorithm; (b) penalty factor (C) in SVM algorithm.

experimental results of SVM algorithm hyperparameter tuning on separate sEMG signals, separate ultrasound signals and fusion of sEMG and ultrasound signals are shown in the Fig.7(b). Analyzing according to Fig.7(b), the optimal penalty factor (C) was 400 as for sEMG, 5 as for ultrasound and 9 as for fusion sEMG-ultrasound.

2) Classification Accuracy: After optimizing the number of principal components and hyperparameters of KNN and SVM, the involved three classifiers were adopted to evaluate the classification accuracy in the involved three types of dataset. The comparison results of sEMG, ultrasound and fusion sEMG-ultrasound recognition accuracies are shown in Fig.8. In terms of recognition accuracy, regardless of the classifier employed, the fusion sEMG-ultrasound feature was always the best, followed by ultrasound feature, and sEMG feature was the worst. Among the three machine learning algorithms, SVM classifier always attains the best recognition rate for these three datasets, followed by LDA classifier, and KNN classifier attains the worst.

B. CNN Extracted Features Method

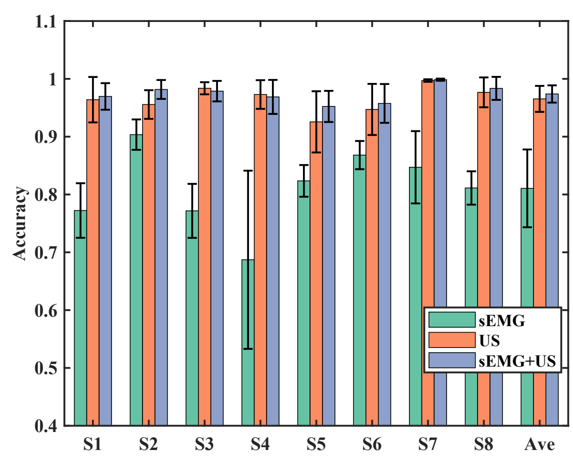

Fig. 9. The classification accuracy of CNN based method across the involved three datasets (separate sEMG, separate ultrasound, fusion sEMG and ultrasound).

The classification accuracy based on convolutional neural network architecture using sEMG signal, ultrasound signal or fusion sEMG-ultrasound signal is present in Fig.9. The average recognition accuracy across all subjects was $81.06 \pm 4.72\%$ in sEMG dataset, $96.53 \pm 2.25\%$ in ultrasound dataset, and $97.38 \pm 1.49\%$ in fusion sEMG-ultrasound dataset.

TABLE I

COMPARISON OF THE AVERAGE RECOGNITION ACCURACY ACROSS ALL SUBJECTS BETWEEN MACHINE LEARNING ALGORITHM (LDA, KNN, SVM) AND CONVOLUTIONAL NEURAL NETWORK.

	sEMG	ultrasound	sEMG-ultrasound
LDA	0.7962	0.9461	0.9601
KNN	0.6654	0.9319	0.9443
SVM	0.7989	0.9520	0.9611
CNN	0.8106	0.9653	0.9738

The average recognition accuracies across all subjects between machine learning algorithm (LDA, KNN, SVM) and convolutional neural network are compared in TABLE I.

The gesture recognition accuracy of ultrasound signals is higher than that of sEMG signals with the adoption of whether the machine learning algorithm or the deep learning algorithm. This is because the ultrasound signal is more sensitive and has better spatial resolution, which makes it able to perform better in discrete gesture recognition. The results is consistent with the conclusion of previous researches.

Whether manually designed features and machine learning algorithms or convolutional neural networks are employed, the classification accuracy of fusion sEMG and ultrasound features is higher than that of separate sEMG features or ultrasound features. It shows that there is supplementary information between sEMG and ultrasound features. The experimental results proves that feature fusion of sEMG and ultrasound signal can take advantage of the complementarity of electrophysiological information and muscle morphology information.

As for feature extraction mathod, when using CNN algorithm to automatically extract features, whether for individual sEMG dataset, individual ultrasound dataset or fusion sEMG-ultrasound dataset, the recognition rate achieved by the CNN algorithm is always higher than the best result achieved by the three machine learning algorithms, which demonstrates that the convolutional neural network can mine more effective features and has better representation ability.

V. CONCLUSION

Through comparison, the ultrasound signal outperforms sEMG signal in classification accuracy and has a significant advantage over sEMG signal in sophisticated gesture recognition tasks.

This paper proposes two feature fusion method. Regardless of the method adopted, fusion sEMG-ultrasound feature always performs better than the individual sEMG or ultrasound features, demonstrating the potential complementarity between the sEMG and ultrasound signal and the feasibility of sEMG-ultrasound feature fusion.

The method of extracting features with CNN has better classification accuracy than the traditional manual feature extraction method, showing the superiority of CNN as a feature extractor. In addition, sEMG and ultrasound signals can share a set of feature extractors based on CNN, which provides inspiration for the feature fusion of multi-modal signals in gesture recognition.

REFERENCES

[1] J. Kim, S. Mastnik, and E. Andr, "Emg-based hand gesture recognition for realtime biosignal interfacing," in *Proceedings of the 2008 International Conference on Intelligent User Interfaces, January 13-16, 2008, Gran Canaria, Canary Islands, Spain*, 2008.

[2] T. Yanagisawa, M. Hirata, Y. Saitoh, H. Kishima, K. Matsushita, T. Goto, R. Fukuma, H. Yokoi, Y. Kamitani, and T. Yoshimine, "Electrocorticographic control of a prosthetic arm in paralyzed patients," *Annals of Neurology*, vol. 71, no. 3, pp. 353–361, 2012.

[3] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *Journal of Rehabilitation Research & Development*, vol. 48, no. 6, p. 643, 2011.

[4] M. Atzori, A. Gijsberts, S. Heynen, A. G. Mittaz-Hager, and H. Mueller, "Building the ninapro database: a resource for the biorobotics community," in *IEEE International Conference on Biomedical Robotics and Biomechatronics*, 2012.

[5] X. Zhai, J. Beth, R. H. M. Chan, and T. Chung, "Self-recalibrating surface emg pattern recognition for neuroprosthesis control based on convolutional neural network," *Frontiers in Neuroscience*, vol. 11, pp. 379–, 2017.

[6] Z. Ding, C. Yang, Z. Tian, C. Yi, Y. Fu, and F. Jiang, "semg-based gesture recognition with convolution neural networks," *Sustainability*.

[7] Y. Hu, Y. Wong, W. Wei, Y. Du, M. Kankanhalli, W. Geng, and H. He, "A novel attention-based hybrid cnn-rnn architecture for semg-based gesture recognition," *Plos One*, vol. 13, no. 10, 2018.

[8] Y. Fang, D. Zhou, K. Li, and H. Liu, "Interface prostheses with classifier-feedback based user training," *IEEE Transactions on Biomedical Engineering*, vol. PP, no. 99, 2016.

[9] Y. Zhou, J. Zeng, K. Li, Y. Fang, and H. Liu, "Voluntary and fes-induced finger movement estimation using muscle deformation features," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 5, pp. 4002–4012, 2020.

[10] Y. Zhou, J. Liu, J. Zeng, K. Li, and H. Liu, "Bio-signal based elbow angle and torque simultaneous prediction during isokinetic contraction," *Science China Technological Sciences*, vol. 62, no. 1, pp. 21–30, 2019.

[11] J. Shi, S. X. Hu, Z. Liu, J. Y. Guo, and Y. P. Zheng, "Recognition of finger flexion from ultrasound image with optical flow: A preliminary study," in *International Conference on Biomedical Engineering & Computer Science*, 2010.

[12] N. Akhlaghi, C. A. Baker, M. Lahlou, H. Zafar, K. G. Murthy, H. S. Rangwala, J. Kosecka, W. M. Joiner, J. J. Pancrazio, and S. Sikdar, "Real-time classification of hand motions using ultrasound imaging of forearm muscles," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1687–1698, 2015.

[13] X. Yang, Z. Chen, N. Hettiarachchi, J. Yan, and H. Liu, "A wearable ultrasound system for sensing muscular morphological deformations," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2019.

[14] X. Yang, X. Sun, D. Zhou, Y. F. Li, and H. Liu, "Towards wearable a-mode ultrasound sensing for real-time finger motion recognition," *IEEE Transactions on Neural Systems & Rehabilitation Engineering*, pp. 1–1, 2018.

[15] W. Xia, Y. Zhou, X. Yang, K. He, and H. Liu, "Towards portable hybrid surface electromyography/a-mode ultrasound sensing for human-machine interface," *IEEE Sensors Journal*, 2019.

[16] W. Guo, X. Sheng, H. Liu, and X. Zhu, "Development of a multi-channel compact-size wireless hybrid semg/nirs sensor system for prosthetic manipulation," *IEEE Sensors Journal*, vol. 16, no. 2, pp. 447–456, 2016.

[17] W. Guo, X. Sheng, H. Liu, and X. Zhu, "Mechanomyography assisted myoeletric sensing for upper-extremity prostheses: A hybrid approach," *IEEE Sensors Journal*, vol. 17, no. 10, pp. 3100–3108, 2017.

[18] W. Bo, R. Wu, K. Zhang, and L. Liu, "A new subtle hand gestures recognition algorithm based on emg and fsr," in *2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 2017.

[19] L. Hargrove, K. Englehart, and B. Hudgins, "The effect of electrode displacements on pattern recognition based myoelectric control," in *Conference International Conference of the IEEE Engineering in Medicine & Biology Society*, 2006.

[20] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Trans on Biomedical Engineering*, vol. 50, no. 7, pp. p.848–854, 2003.