Proceedings of 2020 IEEE
International Conference on Mechatronics and Automation
October 13 - 16, Beijing, China

Convolution Neural Network (CNN)-based Upper Limb Action Recognition

Wenyang Gao[1], Shuxiang Guo[1,2*], Dongdong Bu[1]
[1] *Key Laboratory of Convergence Biomedical Engineering System and Healthcare Technology, The Ministry of Industry and Information Technology, School of Life Science, Beijing Institute of Technology, No.5, Zhongguancun South Street, Haidian District, Beijing 100081, China*
[2] *Faculty of Engineering, Kagawa University, 2217-20 Hayashi-cho, Takamatsu, Kagawa 760-8521, Japan*
E-Mails: {gaowenyang & guoshuxiang & budongdong} @bit.edu.cn;
* Corresponding author

***Abstract* – Surface electromyography (sEMG) can directly reflect human neuromuscular activity, so sEMG is used to track and identify human joint movement in the field of rehabilitation medical engineering. For a large number of sEMG graphs, the results of using traditional classifiers to process the results are unsatisfactory. Using large classic convolution neural network (CNN) to process for a long time will cause a large delay in the control process. In order to make up for the shortcomings and deficiencies, the paper uses lightweight CNN to effectively classify and predict a large number of sEMGs in different parts. Since a lightweight network is used, low latency can be achieved for control effects. Theoretically, a training model with strong real-time performance and high accuracy can be achieved, and it also has a considerable effect in the control process.**

Index Terms - Rehabilitation training. Surface EMG. Convolution neural network.

I. INTRODUCTION

Hemiplegic patients are usually caused by stroke and other brain diseases, which may lead to partial loss of daily living ability of most patients, thereby increasing the burden on families and society [1], [2]. Upper limb function is the most important for many ADLs. Improving the upper limbs' ability after brain injury requires early intensive treatment [3]. In medical practice, medical rehabilitation robots are mainly used to restore the function of the limb movement system [4]. In recent years, surface electromyography (sEMG) as a control source for friendly human-computer interaction in rehabilitation systems has been widely used in motion recognition because of its ability to reflect muscle neuromuscular activity [5]. By preprocessing the collected surface electromyography signals (sEMG) to obtain a feature map that can be input to the network, the convolutional neural network is used to operate the feature map to obtain efficient and accurate classification [22], [24].

Convolutional neural network (CNN) workflow: the pre-processed sEMG signal map is passed to the model, the non-linear unit (activation function) is introduced through the convolution layer, and the number of features to be extracted is reduced by the pooling layer, and Global average pooling, and finally get the result [10], [11].

The convolutional layer recognizes features by setting the size of the convolution kernel, mainly to identify features such as curves and boundaries [6], [7]. Each convolution kernel can be used as a feature recognizer, and the feature map from the previous layer can be calculated by the convolution kernel according to the set step size and filling method for each feature [8]. Doing so reduces network parameters through local connections and parameter sharing [9].

The main functions of the pooling layer are reflected in four aspects: reducing the redundancy of feature information and suppressing interference, increasing the scale invariance and rotation invariance of the model, reducing the computational complexity of the model, and finally preventing overfitting [12], [13]. For a model, the main pressure with many parameters comes from the fully connected layer. In order to reduce the huge parameters brought by the fully connected layer, the convolutional layer before the global average pooling is combined with the convolution layer to replace the fully connected layer [14]. The final result shows that the use of global average pooling has the obvious effect of having fewer training parameters and shorter training time than the fully connected layer, and the recognition accuracy is also higher than that of the fully connected layer, but it is not very obvious [21], [25], [27].

In our previous research, Zhibin Song [15] proposed an improved weighted peaks method to process the filtered sEMG signals from the biceps muscle and adapted linear fitting method to obtain the elbow motion in sagittal plane. Zhenyu Wang [16] utilized the multi-scale entropy and moving-window method to reveal the elbow motion information hidden in the filtered sEMG signals from the biceps muscle. Xuan Song [17], [22] utilized a novel method Ensemble Empirical Mode Decomposition (EEMD) to process the raw sEMG signals and the continuous posture of elbow flexion and extension were recognized based on this method [23]. This paper is to reuse the deep learning convolutional neural network to perform feature extraction and classification tasks on surface EMG signals, to achieve model training under offline conditions, and to find the optimal training model. Thus, the offline model is applied to online control [18], [19].

This paper will obtain the surface electromyography signals of the upper limbs with obvious restoration of health and complete health, to guide the rehabilitation robot to assist the hemiplegic upper limbs for rehabilitation training. Based on previous research, this article mainly did the following work:

First, according to the uneven nature of data collection, it was decided to use different model integration methods for training, and the best performance during training through model integration technology. The effect parameters are saved to obtain the optimal model. The second is pattern recognition of movement. It not only avoids the unreliability of a single accuracy index on unbalanced data, but also solves the real-time control system. In this paper, CNN is used to train and save the optimal model through the offline model training of EMG on the EMG signals of different parts of the upper limb, which can achieve online real-time control to verify the training effect of the model.

II. Method

A. sEMG signal preprocessing

In order to train the model to quickly find the local optimum of the model, this paper normalizes the collected surface myoelectric signal (sEMG) according to the following formulas (1) to (2) to obtain a data form that meets the standard normal distribution. That is, the mean is 0 and the variance is 1. Where μ is the mean and σ is the standard deviation. The purpose of normalizing the data before inputting the data into the model is to make the features in the feature vectors not much different. Speeding up the training speed is to make the entire loop more like a perfect circle rather than an ellipse, and the gradient descent toward the lowest point. Faster, may improve training accuracy.

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \tag{2}$$

B. Unbalanced data model training

There are four processing methods for the unbalanced data and processing methods of the data set, selecting the correct evaluation model index, resampling the training data, integrating different models and redesigning the appropriate cost function. Because the number of patient data sets is much smaller than that of healthy people, this paper adopts the method of integrating different models for this data imbalance data set. At present, there are 5,000 sEMG signal pictures of patients in the data preparation work, and the data set of healthy people reaches more than 40,000, which is a very uneven data set. Because the data set of the healthy person is about 8 times as much as the patient data set, after the data processing method, it is decided to divide it into 8:1 according to the ratio of healthy person data and patient data, as shown in Figure 1. In this process, the patient data set and one of the 8 healthy person data sets after stratified sampling will be used for training to obtain similar 8 different models. Use patient data and each piece of healthy person data are used for model training, and then the best model effect is selected through an integrated method.

Random stratified sampling of the measured data set of healthy persons is divided into 8 data sets that are similar in size to the patient data set, and 8 similar models can be trained by importing the data into the model, and the optimal parameters are selected from these models respectively Parameters as the final model parameters. Among them, the accuracy rate is still used as the evaluation model index, and the choice of resampling data set is limited, so the method of integrating different models is selecte

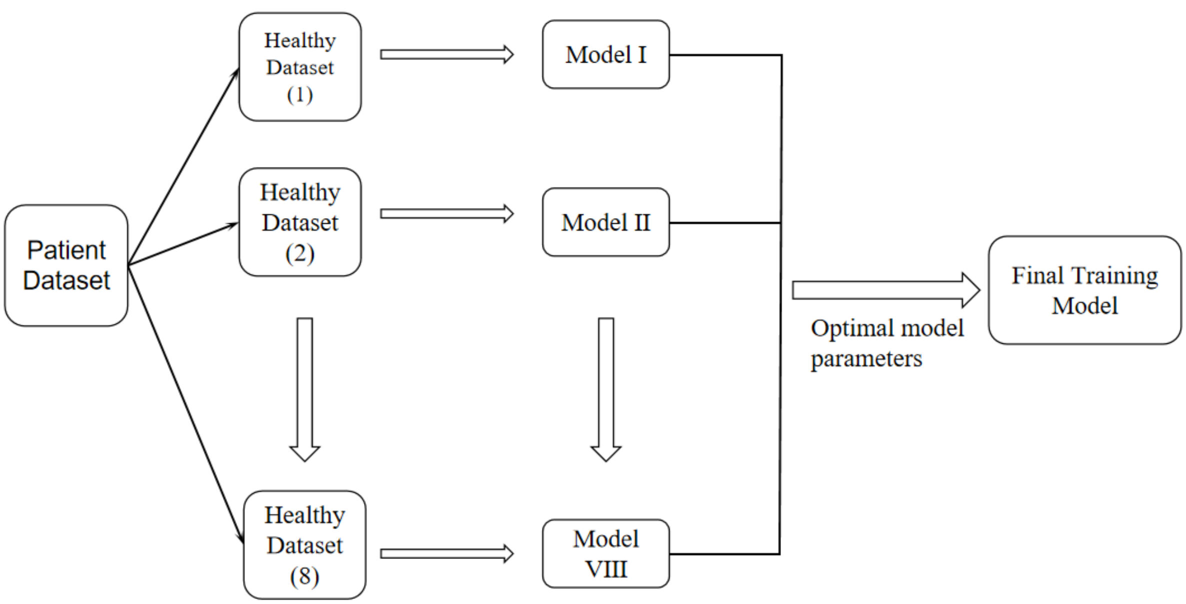
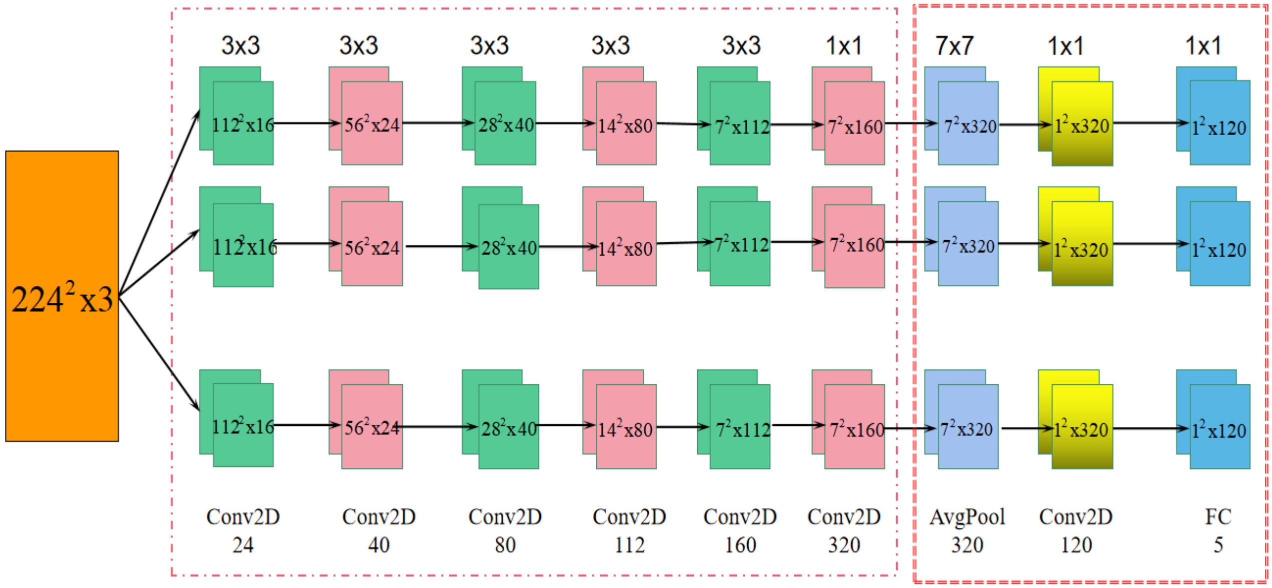


Fig. 1. Integration of different models



a. Convolution layers and Pooling layers

b. AvgPool and FC layers

Fig. 2. Build model structure

C. *Model*

The lightweight CNN used in this article is based on the research of GhostNet by Huawei Noah's Ark Laboratory. The basic parameters of the data set of this article are selected by changing the convolution and the number and step size. Based on the deep learning framework tensorflow, the training model structure is built as follows Figure 2 shows.

The size of the data picture is processed to a size of 224x224x3, and the pixel feature is converted into a feature map of size 7x7x320 through the operation of the convolution layer and the pooling layer. The activation function used is ReLU. Because the network uses Batch Normalization to avoid overfitting, the dropout layer is not used in the back, and the loss function does not use regularization. Then use the global average pooling layer to transform the features into one dimension, use a convolutional layer with a convolution kernel size of 1x1, and finally use the softmax activation function as the final layer to obtain the classification result.

III. Experiments And Results

According to the training results shown in Table 1, it shows the time consumption of the model, the loss value and the accurate removal rate on the training set, and the loss value and the accurate removal rate on the test set. It can be clearly seen that the model is feasible, and the processing time is only a short delay of two and a half minutes to meet the requirements. According to the two accuracy rates, it can be seen that the model has not been over-fitted, so there is no need to add dropout layers and regular terms.

Draw the training results of the second model among the eight models as shown in Figure 3, and compare the loss function and accuracy on the training set and the test set. Due to data reasons, there may be bad models, such as overfitting. This is due to the imbalance of data, so this paper chooses the method of integrating different models to extract the optimal parameters in each training model. The final effect is as shown in Table 1. The delay is short, the accuracy is high, and the result of overfitting does not occur.

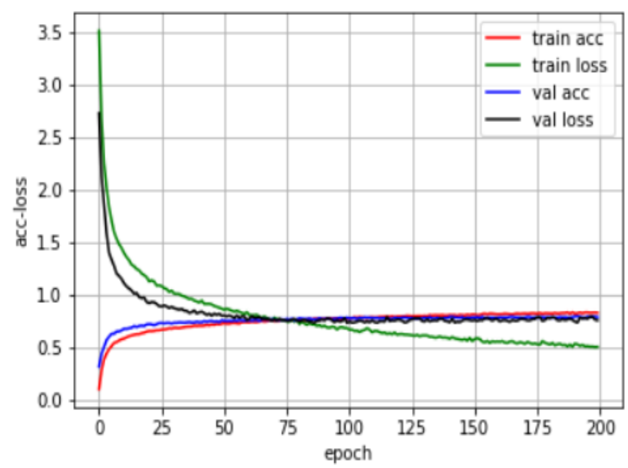


Fig. 3. Model evaluation index

TABLE I
The evaluation indicators of the model

	CNN Model
Train Accuracy	0.8955
Test Accuracy	0.8783
Time	152.0025s
Train Loss	0.16006
Test Loss	0.60017

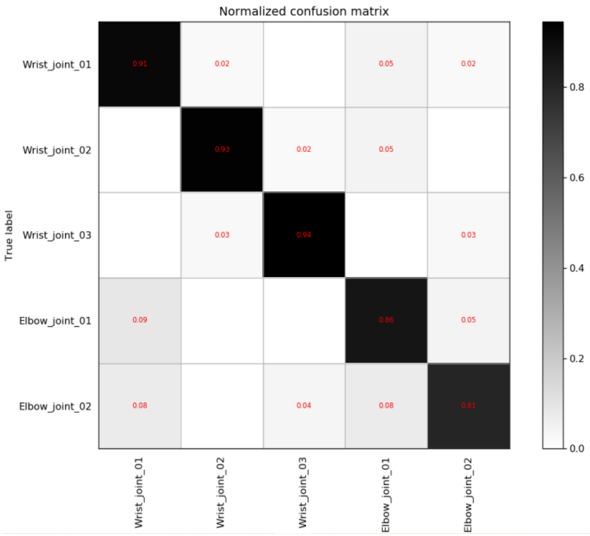


Fig. 4. The confusion matrix for classification prediction

The five movements are arbitrarily named as wrist internal rotation 'Wrist_joint_01', wrist extension 'Wrist_joint_02', wrist joint rotation 'Wrist_joint_03', elbow joint 'Elbow_joint_01' and elbow joint extension 'Elbow_joint_02'. Obtaining the confusion matrix is shown in Figure 4. For the imbalanced data training model, if the model training is directly performed, classification errors may occur, that is to say, a single accuracy rate is not reliable as an evaluation index. According to the confusion matrix, it can be understood that there are erroneous classifications of each category, especially the extension of the elbow joint, which leads to a large amount of information classification errors due to the small amount of data of the patient. It can almost be predicted that the wrist joint rotation 'Probe'. Therefore, it is an effective and feasible method to integrate different models.

According to the confusion matrix, the accuracy of the model is not very high, especially the movement of the elbow joint. The accuracy is greatly affected by the extraction of the selected EMG signal channel, but it can be achieved in the laboratory data. The accuracy rate is over 90%. It may not be very intuitive from the data. Assuming that when a person uses an exoskeleton robot, the number of trainings here is as much as tens of thousands of times or even hundreds of thousands of actions, then a 10% error rate may have thousands or even tens of thousands of misidentified actions. This may be a great hidden danger in actual operation. The focus of this article is to identify actions online, so response time is an important indicator to measure the model.

Timeliness is the innovative point proposed in this paper. After preprocessing the collected signals, the training of the network model is completed in offline mode, so time-consuming has become one of the indispensable indicators of the model. Therefore, the model in this paper is referring to the VGG network model, but the number of convolutional layers and the number of layers are changed, the convolutional layer is reduced, and the number of convolutional kernels is halved. Therefore, the training speed of the VGG model is relatively reduced in time. As shown in Table 3, the training time of the network used in this article is only about one-third of its VGG16. To a large extent, this problem can accelerate timeliness. With the development of deep learning networks, more and more technologies are used on mobile devices. The choice in this article is to design an efficient and lightweight network model.

TABLE II
Model Comparison

	Accuracy	Time
VGG16	0.8750	15min
ResNet	0.8412	11min
Mymodel	0.8783	152.0025s

Through this experiment, this model can be applied to the environment of online real-time application. In comparison, the number of convolution kernels per layer can be reduced to one third of the total number of parameters. From the point of view of the parameter size of the model, the total number of parameters of the model of more than 80M in this paper is approximately one-third compared to the total number of other network parameters. According to the results, the time-consuming situation of this model is not only the shortest and the fastest in response, but also the accuracy of the network model is the highest. The main reason is that Batch Normalization (BN) is used in the network. The selection of a relatively large initial learning rate can speed up the training speed. The addition of the BN layer can effectively avoid overfitting, so that the training effect performance on the test set, the results are not very good. In summary, the experimental scheme proposed in this article is feasible and effective.

The process of this experiment and the particularity of the data source caused the ratio of data to be unsatisfactory. Therefore, the choice of treatment method is to have two options. The first is data enhancement technology, and the second is to select different model evaluation indicators and use multiple models to compare training. This article uses the second method. Although data enhancement technology has a good prospect and scope of application, according to the amount of data and model comparison, the choice is to use multiple models to compare results.

IV. Conclusions

In this paper, an offline model that is effective for motion recognition is trained by the surface electromyography signal (sEMG) between the patient and the healthy person. According to the principle of offline model training, an online real-time control system is obtained. The data processing time combined with offline model training can roughly infer the length of the online delay of the model. This article uses a portable convolutional neural network. The use of Batch Normalization will reduce the training overfitting of the model, and the phenomenon of data imbalance is also improved through the integration of different models. By drawing a confusion matrix to understand how much data has classification errors, a single accuracy rate is unreliable. According to the training results of the model, it is seen that the delay is within the tolerable range. Therefore, the use of portable convolutional neural networks is feasible in pattern recognition and has portability, higher accuracy, and faster training speed. For model training on positive and negative data, choose a better trained model. Use a better performing model and save the training parameters. According to the training results of this experiment, it is clear that the lightweight model network can not only alleviate the hardware requirements of the computer, but also speed up the training speed and reduce the response time. Being able to give subjects' feelings accurately and in real time can provide a more powerful basis for model deployment.

ACKNOWLEDGMENT

This research is partly supported by the National High Tech. Research and Development Program of China (No.2015AA043202)

REFERENCES

[1] S. Zhang, Shu. Guo, Mu. Pang, B. Gao, and P. Guo, "Mechanical Design and Control Method for SEA and VSA-based Exoskeleton Devices for Elbow Joint Rehabilitation," *Neuroscience and Biomedical Engineering*, vol. 2, no. 3, pp. 142-147, 2015.

[2] S. Zhang, S. Guo, B. Gao, H. Hirata, and H. Ishihara, "Design of a Novel Telerehabilitation System with a Force-Sensing Mechanism," *Sensors*, vol. 15, no. 5, pp. 11511-11527, 2015.

[3] B. Gao, H. Ma, S. Guo, H. Xu, and S. Yang, "Design and Evaluation of a 3-degree-of-freedom Upper Limb Rehabilitation Exoskeleton Robot," in proceeding of *2017 IEEE International Conference on Mechatronics and Automation, August 6-9, Takamatsu, Japan*. pp.938-942, 2017.

[4] S. Guo, M. Pang, B. Gao, H. Hirata, and H. Ishihara, "Comparison of sEMG-Based Feature Extraction and Motion Classification Methods for Upper-Limb Movement," *Sensors (Basel)*, vol. 15, no. 4, pp. 9022-9038, 2015.

[5] T. Sharma and K. Veer, "EMG classification using wavelet functions to determine muscle contraction," *Journal of Medical Engineering & Technology*, vol. 40, no. 3, pp. 99-105, 2016.

[6] A. Phinyomark, C. Limsakul, and P. Phukpattaranont. "A novel feature extraction for robust EMG pattern recognition," *Journal of Computing*, vol. 1, no. 1, pp.71–80, 2009.

[7] M. Jiang, R. Wang, J. Wang, and D. Jin, "A Method of Recognizing Finger Motion Using Wavelet Transform of Surface EMG Signal." in Proceeding of *2005 IEEE Engineering in Medicine and Biology*, pp.2672-2674, 2005.

[8] J. Han, Q. Ding, A. Xiong, and X. Zhao. "A State-Space EMG Model for the Estimation of Continuous Joint Movements," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 7, pp. 4267-4275, July, 2015.

[9] G. Saevarsson, J. Sveinsson, and J. Benediktsson, "Wavelet-packet Transformation as a preprocessor of EEG waveforms for classification," in Proceeding of *19th IEEE Annual International Conference on Engineering in Medicine and Biology Society*. pp. 1305–1308, 1997.

[10] X. Chen, X. Zhu, and D. Zhang, "A discriminant bispectrum feature for surface electromyogram signal classification". *Med Eng Phys*, vol. 32, no. 2, pp. 126–135, 2010.

[11] E. Orosco, N. Lopez, and F. Sciascio, "Bispectrum-based features classification for myoelectric control," *Biomedical Signal Processing and Control,* vol. 8, no. 2, pp. 153–168, 2013.

[12] Y. Chen, Y. Zhou, X. Cheng, X. Fan, and Y. Zhang. "Bispectrum-based sEMG multi-domain joint feature extraction for upper limb motion classification," *Mechanical Engineering Science,* vol. 230, no. 2, pp. 248–258, 2016.

[13] Z. Song, Z. Wang, and S. Guo, "SEMG-based Continuous Posture Recognition of Elbow Flexion and Extension in Sagittal Plane," in Proceeding of *2013 IEEE International Conference on Mechatronics and Automation, August 4-7, Takamatsu, Japan*. pp.920-925, 2013.

[14] Z. Wang, S. Guo, B. Gao, and X. Song, "Posture Recognition of Elbow Flexion and Extension Using sEMG Signal Based on Multi-Scale Entropy," in Proceeding of *2014 IEEE International Conference on Mechatronics and Automation, August 3-6, Tianjin, China*. pp.1132-1136, 2014.

[15] Z. Song, Z. Wang, and S. Guo, "Motion Recognition of the Bilateral Upper-limb Rehabilitation using sEMG Based on Ensemble EMD," in Proceeding of *2014 IEEE International Conference on Mechatronics and Automation, August 3-6, Tianjin, China*. pp.1637-1642, 2014.

[16] Y. Jiang, X. Chen, J. Tao, and C. Zhang, "Numerically simulating non-Gaussian random processes with specified PSD, skewness and kurtosis," *Journal of System Simulation*, vol. 18, no. 5, pp. 1127–1130, 2006.

[17] S. Yu and M. Lee, "Bispectral analysis and genetic algorithm for congestive heart failure recognition based on heart rate variability," *Computers in Biology and Medicine,* vol. 42, no. 8, pp. 816–825, 2012.

[18] Y. Zhou, J. Chen, G. Dong W. Xiao, and Z. Wang, "Application of the horizontal slice of cyclic bispectrum in rolling element bearings diagnosis," *Mechanical Systems and Signal Processing*, vol. 26, no. 6, pp. 229–243, 2012.

[19] R. Martis, U. Acharya, and L. Min, "ECG beat classification using PCA, LDA, ICA and discrete wavelet transform," *Biomedical Signal Processing and Control*, vol.8, no. 5, pp. 437–448, 2013.

[20] S. Zhang, S. Guo, B. Gao, Q. Huang, M. Pang, H. Hirata, and H. Ishihara, "Muscle Strength Assessment System using sEMG-Based Force prediction Method for Wrist Joint," *Journal of Medical and Biological Engineering*, vol. 36, no. 1, pp. 121-131, 2015.

[21] Ziyi Yang, Shuxiang Guo, Yi Liu, Hideyuki Hirata, Takashi Tamiya, "An Intention-based Online Bilateral Training System for Upper Limb Motor Rehabilitation", Microsystem Technologies, DOI:10.1007/s00542-020-04939-x, 2020.

[22] Songyuan Zhang, Qiang Fu, Shuxiang Guo, Yili Fu, "A Telepresence System for Therapist-in-the-loop Training for the Elbow Joint Rehabilitation", Applied Sciences, Vol.9, No.8, DOI: 10.3390/app9081710, 2019.

[23] Songyuan Zhang, Qiang Fu, Shuxiang Guo, Yili Fu, "Coordinative Motion-based Bilateral Rehabilitation Training System with Exoskeleton and Haptic Devices for Biomedical Application", Micromachines, Vol.10, No.1, DOI:10.3390/mi10010008, 2018.

[24] Qiang Fu, Songyuan Zhang, Shuxiang Guo, Jian Guo, "Performance Evaluation of a Magnetically Actuated Capsule Microrobotic System for Medical Applications", Micromachines, Vol.9, No.12, DOI: HYPERLINK "https://dx.doi.org/10.3390/mi9120641"10.3390/mi9120641, 2018.

[25] Yi Liu, Shuxiang Guo, Hideyuki Hirata, Hidenori Ishihara, Takashi Tamiya, "Development of a powered variable-stiffness exoskeleton device for elbow rehabilitation", Biomedical Microdevices, Vol.20, No.3, DOI: 10.1007/s10544-018-0312-6, 2018.

[26] Songyuan Zhang, Shuxiang Guo, Yili Fu, Luc Boulardot , Qiang Huang, Hideyuki Hirata and Hidenori Ishihara, "Integrating Compliant Actuator and Torque Limiter Mechanism for Safe Home-Based Upper-Limb Rehabilitation Device Design", Journal of Medical and Biological Engineering, Vol.37, No.3, pp.357-364, DOI: 10.1007/s40846-017-0228-2, 2017.

[27] J. Stendahl, and A. Sinusas, "Nanoparticles for cardiovascular imaging and therapeutic delivery, part 1: compositions and features," *Journal of Nuclear Medicine Official Publication Society of Nuclear Medicine*, Vol. 56, No. 10, pp. 1469-1475, 2015.