

A User Independent Method for Identifying Hand Gestures with sEMG

Hitoshi Tamura^a, Kazuki Itou and Yasushi Kambayashi
Department of Robotics, Nippon Institute of Technology, Saitama, Japan

Keywords: Semg, Hand Gesture, Deep Learning, LSTM, Data Augmentation.

Abstract: We propose a method to determine hand gestures using sEMG (surface Electromyogram) measured from the forearm. The detection method uses the LSTM (Long Short Term Memory) model of RNN (Recurrent Neural Network). Although the conventional method requires the learning data of the user, this is a method that an unspecified number of users can use immediately by enhancing the data. We have confirmed that the accuracy does not change even if the mounting position of the sensor is shifted. We have shown the effectiveness of the data enhancement by numerical experiments.

1 INTRODUCTION

People use hand gestures as a means of communication. Even though we use speech and facial expressions as the main means of communication, we use hand gestures as part of natural body language. In addition, we have organized the hand gestures as sign language and have been using it for conversation. Therefore, it is natural to think that we may employ the hand gestures as an interface to electronic devices and robots. In order to have a machine recognize hand gestures, we generally use two means. One is a method of using computer vision technique. That is recognizing a hand shape and its operation using imaging devices such as cameras and depth sensors.


The other method is to acquire the shapes of the fingers and three-dimensional acceleration through sensors. In order to do so, we need to attach sensors to the fingers. In the former, external sensors such as cameras and depth sensors are required, and restrictions such as shooting range and effective distance are often imposed on the position of the sensor. In the latter, the human movements are restricted. In addition, the users have to endure from wearing such devices. Both methods need to be adjusted according to the shooting environment and the individual when using them.

Many researchers have proposed variations of both methods. They are called classical methods, and are summarized in the reference (Mitra and Acharya,

2007). Information obtained from sensors and cameras is used to classify gestures. We can utilize various classification methods such as HMM (Hidden Markov Model), FSM (Finite State Machine) and PCA (Principal Component Analysis).

Recently, we have witnessed a remarkable development of machine learning methods. Since the machine learning methods have dramatically improved the performance of the classifiers, many identification problems have been solved. Especially, many researchers have made neural networks perform machine-learning through the measured sEMG (surface Electromyogram) of forearms to identify hand gestures

In this paper, we propose a method that identifies hand gestures by classifying sEMG obtained from forearms using commercially available sensors and deep learning method. The sensors used are Myo Gesture Control Armbands (hereinafter Myo) manufactured by Talmic Labs (now North). They are easy to attach and detach. We made eight sEMG sensors connected like a bracelet so that they are easily detachable from the forearm (Figure 1). The signals acquired from the eight sEMG are transmitted to a control device, i.e. a PC via the Bluetooth connection. Each sensor equips one 3-axis accelerometer and one 3-axis gyroscope. The weight of this sensor is 93g and the thickness is 11 mm. Therefore we can expect the users to feel little discomfort when they wear. Myo is suitable for direct operation in VR and AR.

^a <https://orcid.org/0000-0002-7836-3042>

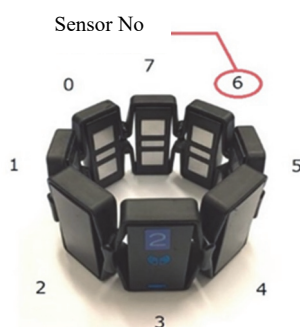


Figure 1: Myo Gesture Control Armband.

By using such simple sensors, we do not need to install cameras, which are required in conventional methods, nor attaching devices to fingers such as data globes, which require wire connections and adjustments. Furthermore, since mounting sEMG devices imposes little discomfort to the users, we can provide more natural interfaces.

For hand gesture analysis, several conventional machine learning methods have been proposed for learning sEMG. They are methods using decision trees and HMM for sign language recognition (Zhang et al., 2011), methods using decision trees and k-NN (k-Nearest Neighbor) for hand gesture analysis (Lian et al., 2017), methods using PCA for prosthetic control (Matrone et al., 2011), methods using HMM and SVM (Support Vector Machine) (Rossi et al., 2015), and methods using an application of ANN (Artificial Neural Network) to hand gesture analysis (Liu et al., 2017).

The methods of classifying the pattern of sEMG by machine learning are roughly divided into two categories: one is dealing with static gestures and the other is that of including dynamic gestures. For dealing with only static gestures, it is sufficient to analyze a few snapshots for some moments of sEMG. In order to analyze general dynamic gestures, however, it is necessary to obtain time series gesture data.

In order to classify time series data of sEMG for dynamic hand gestures, we use RNN (recurrent neural network). RNN is suitable for time series data. One particular RNN is especially suitable for time series data. That is LSTM (long short term memory) model. It is an extended version of RNN.

The idea of analyzing sEMG by RNN is not new. It has been employed in the field of biomedical engineering and robotics since 1990's. It has been used to estimate the angles of joints in a human body from sEMG, and to calculate motor control parameters that control robots, electric prosthetic feet, power assist suits (Koike et al., 1993; Koike et al., 1994; Koike et al., 1995; Cheron et al., 1996; Cheron

et al., 2003).

Applying LSTM to sEMG time series data to classify gestures are found in (Wu et al., 2018; Samadani et al., 2018; Quivira et al., 2018). According to those experiments, LSTM improves the accuracy of the classification.

The problem is that when using a simple sEMG sensor such as Myo to identify the hand gesture, a slight deviation of attaching the device greatly affects the acquired values of sEMG. It is difficult to measure the complex movement of the forearm muscles that are complex three-dimensional shapes with a sensor that can measure only the muscle potential of the body surface. Conventionally, this problem is avoided by providing a large amount of data for machine-learning. However, this method requires a large amount of data for each user, and the learned classifier is effective only for that particular user. It is difficult to make classifiers ready for unspecified number of users.

There are many proposals to classify sEMG by a classifier that is built by machine learning. However, most of them are tailored for a specific user. Because they are built from the user's own learning data. A user has to provide a large amount of his or her own data for the machine learning. There is no known attempt of trying to build a learned classifier for unspecified number of users.

In this paper, we report our attempt to develop a hand gesture classifier that can be applied to an unspecified number of people by effectively augmenting several sEMG data.

2 TARGET GESTURES AND DETERMINATION METHODS

2.1 Types of Target Gestures

Figure 2 shows the target gestures. We classified six types of gestures (weakness, paper, lightly grasping, strongly grasping, pointing finger, and scissors).

We measured each gesture, and took for each four seconds. Since we wanted to have practical setting, we did not exclude the duration time sEMG being stabilized. We started to measure sEMG for four seconds when the user started to perform each gesture.



Figure 2: Hand Gestures to measure.

Output Gate, and Forget Gate) in the hidden layer. The hidden layer is called LSTM Block. Owing to this hidden layer, the LSTM exhibits high discriminant performance for time series data. Figure 3 shows the structure of LSTM, and Figure 4 shows the network that was actually used for learning.

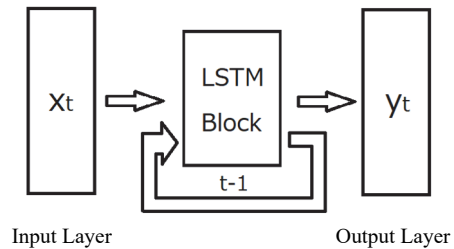


Figure 3: LSTM Structure.

2.2 Configuring LSTM Networks

We used SONY's Neural Network Console (NNC) as an integrated development environment for deep learning. We have fine-tuned the LSTM included in the NNC sample and used for learning. The LSTM that we used for learning has three gates (Input Gate,

2.3 Method for Obtaining Learning Data

In this study, we constructed a system that records sEMG data obtained from Myo and stores in a CSV file for every 10ms (Figure 5). The digital values of

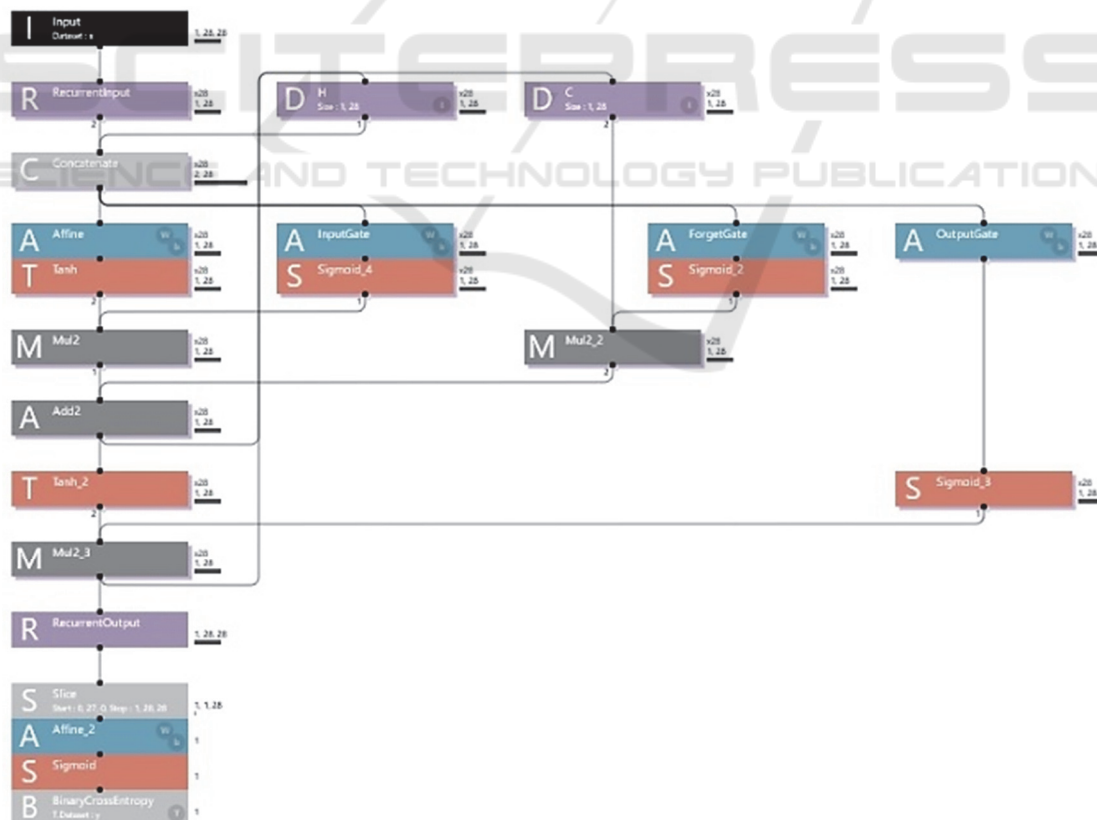


Figure 4: Network Used for Learning.

the acquired sEMG are output in the range of -127 to 127. In order to learn in the NNC, however, we had to convert the data into the normalized data ranged -1 to 1 before outputting to the CSV file. Figure 6 shows an example of the output CSV file .



Figure 5: Measurement system.

	A	B	C	D	E	F	G	H
1	-0.03125	-0.03125	0.007813	-0.02344	-0.01563	0	-0.02344	0
2	-0.02344	0.03125	0.007813	-0.02344	-0.01563	0	-0.02344	0
3	0.007813	0.039063	-0.02344	-0.02344	-0.00781	-0.00781	0.015625	0
4	0	0.023438	-0.00781	-0.01563	0	-0.00781	0.015625	0
5	-0.04688	-0.0625	-0.04688	-0.03125	-0.01563	-0.00781	-0.02344	-0.03125
6	-0.01563	0.03125	0.015625	-0.01563	0.007813	0	-0.04688	-0.00781
7	-0.01563	0.015625	0.015625	0.03125	0.039063	-0.02344	0.046875	0.101563
8	0.039063	0.015625	0.015625	0.03125	0.039063	-0.02344	0.046875	0.101563

Figure 6: Example of a measured CSV file.

2.4 Data Augmentation

We added some random noise data range of -0.15 to 0.15 to each measured sEMG datum. Further augmented data smoothed by taking the average value for every 20ms. We then generated ten new data with random noise from one measurement datum and added to the original data. We made learning be done with dataset eleven times as the original data.

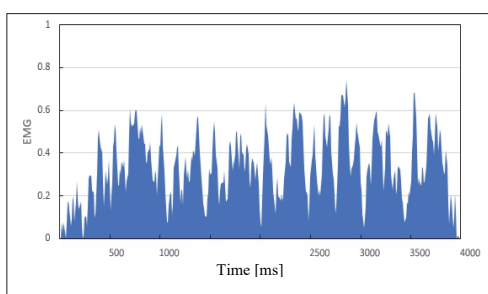


Figure 7: Measured sEMG data.

Figure 7 and 8 show an example of the data measured by sensor No. 0 and the corresponding newly generated data by adding noise respectively.

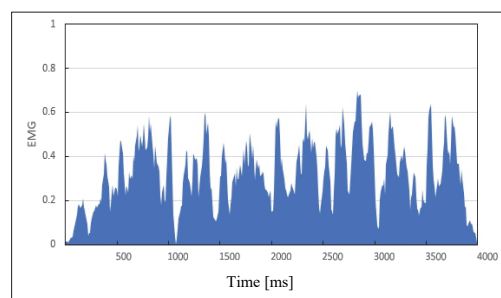


Figure 8: sEMG data with augmented processing.

3 EXPERIMENT

3.1 Accuracy of Reattaching Myo

The author measured his own learning data (2,880). When we measured Myo continuously for learning data and evaluation data without attaching and detaching, the accuracy was 96.38%. On the other hand, when we measured Myo for learning data, detached Myo, and then measured the evaluation data by attaching the same position as the previous time as much as possible, the accuracy of the evaluation data, which were determined by the previous learned parameters, was 35.00%. We can observe that the accuracy decreased significantly by the deviation at the time of mounting.

We measured the learning and evaluation data by re-wearing Myo at seven angles deliberately shifted by 15 degrees for the deviation. We made the machine learn with a total of 22,680 data. The evaluation data were data measured after re-wearing Myo. As a result, we found 93.3% accuracy even when there was a deviation of re-wearing Myo. From this, it can be said that it is possible to prevent the decrease of the accuracy due to deviation at the time of mounting, by mounting it in several places.

Table 1 shows the evaluation results when the learning data was measured by wearing Myo at seven angles. The accuracy shown in Table 1 is the percentage of correctly determined in all inferences. The precision is the rate of correctly determined and estimated to be true. The recall is the percentage of estimated to be true if the data is true. The F-measures is the harmonic mean of the precision and the recall.

Table 1: Evaluation results of learning.

Accuracy	93.3%
Precision	94.2%
Recall	93.3%
F-Measures	93.3%

3.2 Gesture Classification on Different Subject's sEMG

We have investigated whether an unspecified number of users can use the current learned parameters. We have employed three new subjects, and repeated the measurement 60 times. We have used the previously learned parameters described above, and determined with the data of three people. As a result, the average of the correct answer rate of the three people significantly decreased to 43.89%. Table 2 shows the discriminant results.

Table 2: Determination results for each subject.

Gesture Type	Recall of Subject A	Recall of Subject B	Recall of Subject C
Paper	40.0%	0.0%	70.0%
lightly Grasping	20.0%	20.0%	0.0%
Strongly Grasping	80.0%	50.0%	80.0%
Pointing Finger	80.0%	60.0%	100%
Scissors	60.0%	0.0%	0.0%
weakness	70%	60.0%	0.0%
Average	58.3%	31.7%	41.7%

When examining the gestures individually, the "strongly grasped" and "pointing finger" gestures provided a high accuracy regardless of the individuals, but "lightly grasping" and "scissors" gestures hardly provided any accuracy. Figures 9 and 10 show examples of sEMG measurement data for the "lightly grasped" and "scissor" gestures. We can observe that the signals created by "lightly grasping" and "scissors" actions display variety of wave forms. They clearly differ from each other in the way of applying muscle power. It seems that each different individual applies his or her muscle power for "lightly grasping" and "scissors" in a quite unique way.

3.3 Results of Data Augmentation with Random Noise

In order to find out whether the accuracy can be improved by the data augmentation, we have acquired additional data. The subjects were asked to wear Myo at three angles and we measured 3,240 data. In the experiments, we added some random noise data to these collected data, and augment the number of data eleven times as many as the original ones, and made perform learning with 35,640 data. We measured another set of 3,240 evaluation data in the same manner as the learning data. As a result of learning

only the measured data, the accuracy was 75.03%. The accuracy of learning that augmented the data was 78.48%. In other words, the accuracy improved by 3.45%.

We observed a little improvement of the accuracy. Even though for some gesture, we observed a case where learning with only measurement data showed higher accuracy than learning with augmented data, in general, learning with augmented data displays better accuracy. Table 3 shows the discriminant accuracy of each gesture of the classifier learned only by the measurement data and the classifier learned using the augmented data.

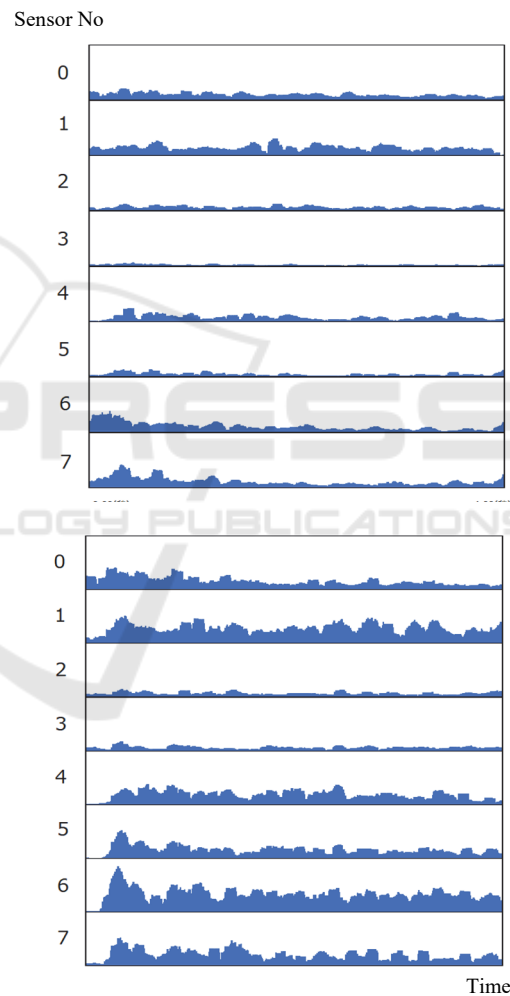


Figure 9: Lightly Grasp (Upper: Learning Data, Lower: Subject A).

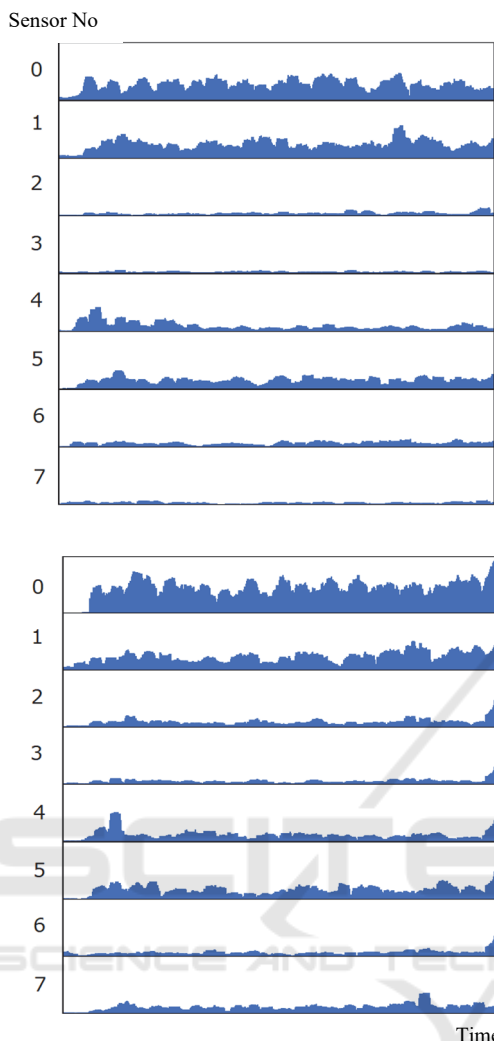


Figure 10: Scissors (Upper: Learning Data, Lower: Subject A).

Table 3: Changes in detection rates due to data augmentation.

Gesture Type	Recall of Measured Data	Recall of Augmented Data
Paper	61.7%	50.0%
lightly Grasping	82.6%	77.0%
Strongly Grasping	94.1%	96.7%
Pointing Finger	85.9%	64.8%
Scissors	31.5%	82.4%
weakness	94.4%	100.0%

Table 4 and Table 5 show the evaluation results by each classifier.

Table 4: Evaluation Results for Measurement Data only.

Accuracy	75.0%
Precision	77.3%
Recall	75.0%
F-Measures	74.0%

Table 5: Evaluation Results with augmented Data.

Accuracy	78.5%
Precision	80.4%
Recall	78.5%
F-Measures	78.0%

4 DISCUSSION

By attaching Myo and changing the angle several times, we could improve the accuracy even if the subjects re-wear Myo. The reason for this phenomenon might be the leaning does not only depends on the value of each sEMG sensor, but also it depends on the numerical balance of the eight sEMG sensors. By learning the data with unfixed angle, we could avoid the over-fit that depends on a specific sensor.

We have found that different individuals provide different output values of sEMG even with the same gesture. In order to avoid overfitting individual-dependent features, it is necessary to measure data from a large number of people. The need to collect data from many people is clear from the result that the accuracy decreased to 43.89% when we applied one specific person's data to others.

On the other hand, the method of changing the angle while measuring requires a large number of repetitions of measurement, thus requires an extremely long measurement time. In order to measure data efficiently, data expansion is essential.

With parameters learned by a specific individual's data, the accuracy becomes very low as determining the gesture of another person. One of the reason is that the way of the muscle power applied to the fingers of each person is different even in the same gesture movement. For example, in the case of the "scissors" gesture, some people don't put any muscle force on their thumbs, and some other people put their thumb and ring finger on top of each other. Even performing the same gesture, there are different patterns in which applying the muscle force. It is difficult to determine the gesture in such cases.

The second reason is that the effect of individual differences in the muscle strength of the entire hand. Even with the same gesture, the condition of applying muscle force to the whole hand is different. Although

sEMG may be able to solve this by normalizing the width between the maximum value and minimum value of the measured value, it may be difficult to distinguish between the state of straining muscle and the state of relax. Therefore, to grasp the force level of each subject in advance, it is necessary to match the process of some criteria. For example, we can divide the power levels into three stages of weak, medium, and strong, and then instruct the subjects to gesture at the level of "medium".

In machine learning, by collecting and learning data from a large number of subjects, it should be possible to generate a classifier that is not affected by individual differences, such as strength of force and differences in finger usage. However, it is too expensive to collect a large amount of data that needs to be physically measured. Therefore, data augmentation is also important in this perspective.

In this data augmentation, we have added some noise data directly to the sEMG sensor measurement data. However, we are planning to add random noise only to the fine features that is maintaining the characteristics of frequency spectrum envelope just as the analysis method of the audio signal. With this new data extension method, it may be possible to generate artificial data with similar characteristics to the measurement data. Although not included in this paper, the preliminary experiments suggest that a new data extension method is effective.

5 CONCLUSION

Hand gestures are not only providing a means of communication between people, but also attracting attention as a method for operating electronic devices and robots. Conventional recognition method using computer vision requires camera and method using sensors requires wearing glove-type devices.

In recent years, the performance of the classifier by machine learning method such as deep learning has been improved. Therefore, there are many studies that try to improve the discriminant accuracy by learning sEMG of hand gestures. In this study, we measured sEMG using the armband type device Myo, which is easy to attach and detach, and learning by the network of LSTM model of RNN, and experimented with the method of determining the hand gestures for an unspecified number of subjects.

We have performed the following experiments.

1. Discriminant accuracy by wearing Myo with deviation.

2. Discriminant accuracy in classifiers learned by data measured by deliberately shifting angles.
3. Expansion of learning data by random noise.

The summary of the experimental results are as follows.

1. By learning with the data measured by wearing Myo from multiple angles, the inaccuracy due to the wearing deviation is reduced, and robustness is improved.
2. As the data extension, the improvement of the discriminant accuracy can be expected by adding noise.

As a future direction, we will try to reduce the influence of individual muscle force by measuring at the strength level of weak, medium, and strong. In regard to the data augmentation, we are planning to develop an interface that can be used by anyone with minimal adjustment by trying a method to generate similar to artificial data that maintains the characteristics of spectral envelope.

REFERENCES

- S. Mitra and T. Acharya, 2007. Gesture recognition: A survey, *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, Vol. 37, No. 3, pp. 311–324.
- X.Zhang, X.Chen, Y.Li, V.Lantz, K.Wang, and J. Yang, 2011. A framework for hand gesture recognition based on accelerometer and EMG sensors. *IEEE Trans. Syst., Man, Cybern. A, Syst. and Humans*, Vol. 41, No. 6, p.1064-1076.
- Kuang-Yow Lian, Chun-CHieh Chiu, Yong-Jie Hong and Wen-Tsai Sung.. 2017. Wearable Armband for Real Time Hand Gesture Recognition. *IEEE International Conf. Syst., Man, Cybern.* pp. 2992-2995.
- G. Matrone, C. Ciprian, M. C. Carrozza and G. Magene, 2011. Two-Channel Real-Time EMG control of a Dexterous Hand Prosthesis. *Proceedings of the 5th International IEEE EMBS Conf. on Neural Eng.* pp. 554-557.
- M. Rossi, S. Benatti, E. Farella, and L. Benini, Hybrid EMG classifier based on HMM and SVM for hand gesture recognition in prosthetics, 2015. *Proc. Int. Conf. Ind. Technol.*, pp. 1700–1705.
- Xilin Liu, Jacob Sacks, Milin Zhang, Andrew G. Richardson, Timothy H. Lucas, and Jan Van der Spiegel, 2017. The Virtual Trackpad: An Electromyography-Based, Wireless, Real-Time, Low-Power, Embedded Hand-Gesture-Recognition System Using an Event-Driven Artificial Neural Network, *IEEE Trans. Circuits and Sys.*, Vol.64, No.11, pp. 1257-1261.
- Koike Y, Honda k, Hirayama R, Eric V and Kawato M., 1993. Estimation of Isometric Tongues from surface electromyography using a neural network Models.

- IEICE Trans. Inf. & Syst. (Japanese Edition)*, Vol. J76-D2, No.6, pp. 1270-1279.
- Koike Y. and Kawato M., 1994. Trajectory formation surface EMG signals using a neural network model, *IEICE Trans. Inf. & Syst. (Japanese Edition)*, Vol. J77-D2, No.1, pp. 193-203.
- Koike Y, Kawato M., 1995. Estimation of dynamic joint torques and trajectory formation from surface electromyography signals using a neural network model. *Biol. Cybern.* Vol.73, No.4, pp. 291-300.
- Cheron G, Draye JP, Bourgeois M, Libert G., 1996. A dynamic neural network identification of electromyography and arm trajectory relationship during complex movements, *IEEE Trans. Biomed. Eng.* Vol.43, No.5, pp.552-558.
- Cheron G, Leurs F, Bengoetxea A, Draye JP, Destrée M, Dan B., 2003. A dynamic recurrent neural network for multiple muscles electromyographic mapping to elevation angles of the lower limb in human locomotion, *Journal of Neuroscience Methods*, Vol.129, No.2, pp.95-104.
- Yuheng Wu, Bin Zheng and Yongting Zhao, 2018. Dynamic Gesture Recognition Based on LSTM-CNN, *Chinese Automation Congress*, 2018, pp.2446-2450.
- Ali Samadani, 2018. Gated Recurrent Neural Networks for EMG-Based Hand Gesture Classification: A Comparative Study, *40th Annual International Conf. of the IEEE Engineering in Medicine and Biology Society*, pp.1094-1097.
- Fernando Quivira, Toshiaki Koike-Akino, 2018. Ye Wang and Deniz Erdogmus, Translating sEMG Signals to Continuous Hand Poses using Recurrent Neural Networks, *IEEE EMBS International Conf. on Biomedical & Health Informatics*, pp. 166-169.
- G.-C. Luh, Y.-H. Ma, C.-J. Yen, and H.-A. Lin, 2016. Muscle-gesture robot hand control based on semg signals with wavelet transform features and neural network classifier, in *Machine Learning and Cybernetics (ICMLC)*, vol. 2. IEEE, pp. 627-632.
- M. Haris, P. Chakraborty, and B. V. Rao, 2015. Emg signal based finger movement recognition for prosthetic hand control, in *Communication, Control and Intelligent Systems (CCIS)*, IEEE, 2015, pp. 194-198.
- M. E. Benalcazar, A. G. Jaramillo, J. A. Zea, A. P'aez, and V. H. 'Andaluz, 2017. Hand gesture recognition using machine learning and the Myo armband, *25th European Signal Processing Conference (EUSIPCO)*, pp.1040-1044.
- Guan-Chun, L., Heng-An, L., Yi-Hsiang, M., ChienJung, Y., 2015. Intuitive Muscle-Gesture Based Robot Navigation Control Using Wearable Gesture Armband. In: *2015 International Conference on Machine Learning and Cybernetics (ICMLC)*, vol. 1, pp. 389-395. IEEE Conference Publications.
- M. Hioki and H. Kawasaki, 2012. Estimation of finger joint angles from sEMG using a neural network including time delay factor and recurrent structure, *Int Scholarly Res Netw (ISRN) Rehabil*, vol. 2012.
- F. E. R. Mattioli, E. A. Lamounier, A. Cardoso, A. B. Soares, and A. O. Andrade, 2011. Classification of EMG signals using artificial neural networks for virtual hand prosthesis control, in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 7254-7257.
- C. Choi and J. Kim, 2007. A Real-time EMG-based Assistive Computer Interface for the Upper Limb Disabled, *Proceedings of the 2007 IEEE 10th International Conf. on Rehabilitation Robotics*, pp. 459-462.
- M. Arveti, G. Gini, and M. Folgheraiter, 2007. Classification of EMG signals through wavelet analysis and neural networks for controlling an active hand prosthesis, *Proceedings of the 2007 IEEE 10th International Conference on Rehabilitation Robotics*, pp. 531-536.
- V. P. Singh and D. K. Kumar, 2008. Classification of low-level finger contraction from single channel Surface EMG, *30th Annual International IEEE EMBS Conference Vancouver*, pp. 2900-2903.
- H. Kawashima, N. Tsujiuchi, and T. Koizumi, 2008. Hand Motion Discrimination by EMG Signals without Incorrect Discriminations that Elbow motions Cause, *30th Annual International IEEE EMBS Conference Vancouver*, pp. 2103-2107.
- J. U. Chu, I. Moon, and M. S. Mun, A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand, *IEEE Transactions on Biomedical Engineering*, Vol. 53, pp. 2232-2239, July. 2006.
- H. Huang and C. Chen, 1999. Development of a Myoelectric Discrimination System for a Multi Degree Prosthetic Hand, *Proceeding of the 1999 IEEE, International Conference on Robotics & Automation*, pp. 2392-2397.
- K. Englehart, B. Hudgins, and P. A. Parker, 2001. A wavelet-based continuous classification scheme for multifunction myoelectric control, *IEEE Trans. Biomed. Eng.*, vol. 48, no. 3, pp. 302-311.
- D. Peleg, E. Braiman, E. Yom-Tov, and G. F. Inbar, 2002. Classification of finger activation for use in a robotic prosthesis arm, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 10, no. 4, pp. 290-293.
- Y. Yazama, M. Fukumi, Y. Mitsukura, and N. Akamatsu, 2003. Feature analysis for the EMG signals based on the class distance, in *Proc. 2003 IEEE Int. Symp. Computational Intell. Robotics Autom.*, vol. 2, pp. 860-863.
- Frederic Kerber, Michael Puhl and Antonio Krüger, 2017. User-Independent Real-Time Hand Gesture Recognition Based on Surface Electromyography, *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, No.36, pp.1-7.