Hand Motion Estimation using Super-Resolution of Multipoint Surface Electromyogram by Deep Learning

Keigo FUKUSHIMA, Yoshiaki YASUMURA
Graduate School of Engineering and Science
Shibaura Institute of Technology
Tokyo, Japan

***Abstract*—This paper proposes a method for hand motion estimation using super-resolution of multipoint surface electromyogram for prosthetic hands. In general, obtaining more EMGs (ElectroMyoGraphy) improves the accuracy of hand motion estimation, but it is costly and hard to use. Therefore, this method improves the accuracy of hand motion estimation by estimating a large number of EMG signals from a small number of EMG signals using super-resolution. This super-resolution is achieved by learning the relationship between few and many myoelectric signals using a deep neural network. Then, hand motions are estimated from the high-resolution signal using a deep neural network. Experiments using actual EMG signals show that the proposed method improves the accuracy of hand motion estimation.**

***Keywords*—Hand motion estimation; super-resolution; deep neural network; prosthetic hand; electromyography**

I. Introduction

Today, many people utilize prosthetic hands in their daily lives due to accidents or congenital disabilities. Myoelectric prostheses record the electrical signals on the skin surface when muscles are moved, and the prosthetic hand moves based on these signals. This allows for intuitive movement with few restrictions on posture.

As described above, myoelectric prostheses are highly useful, but they have some problems. The prosthetic hand can be manipulated by the wearer's intention, though it is difficult for the wearer to perform all of the activities of daily living. To enable more complex hand motions, the accuracy of electromyographic analysis requires improvement. For this purpose, multipoint surface electromyography is useful [1], [2]. Multipoint surface electromyography is a measurement method of muscle movement using a grid of electrodes on the skin surface. This allows us to measure more detailed temporal and spatial myoelectric information.

Jiangcheng et al. placed 128 electrode channels on the upper arm and classified 27 gestures with 95.3% accuracy [3]. Weijie et al. also showed that a larger number of sensors results in a higher percentage of correct gestures [4]. However, increasing the number of electrodes makes mounting more difficult and electrode costs higher. The number of electrodes needs to be reduced to make it easier to use the prosthetic hand.

In order to operate a prosthetic hand with fewer electrodes, we propose a method to improve prosthetic hand motion by increasing myoelectric data with super-resolution technology. The objective of this study is to accurately estimate hand motion from fewer EMG signals. This method estimates multi-point surface electromyograms from fewer electrodes and operates the prosthetic hand more accurately based on them. For this estimation, a convolutional neural network, which is widely used in image recognition, is used for super-resolution of electromyograms and estimation of hand motions. Super-resolution is a technique to obtain a high-resolution output signal from a low-resolution input signal. This method uses deep learning to learn the relationship between low-resolution and high-resolution signals to create a higher resolution.

II. Related Works

Since many people use prosthetic hands in their daily lives, many studies have tried to create better prosthetic hands [5]-[9]. A myoelectric prosthetic hand uses the electrical signals generated on the surface of the skin during muscle movement to operate a robotic hand. Myoelectric prosthetics are more useful than other prosthetics, but there are some problems. Myoelectric prosthetics allow some movements, but complex movements are difficult. To enable complex movements, it is necessary to measure many myoelectric signals and estimate movements from these signals.

Hudgins et al. use the multilayer perceptron to identify four forearm movements from myoelectric signals [1]. Ali et al. used a convolutional neural network to identify 10 hand motions from myoelectric signals acquired at 8000 Hz using six-channel electrodes and 0.15 seconds of myoelectric signals as input [2]. This paper shows that with the appropriate hyper parameters and network structure, it is possible to estimate with more than 80% accuracy for all movements. However, the accuracy needs to be further improved for actual use of the prosthetic hand.

Jiangcheng et al. placed 128 electrode channels on the upper arm and classified 27 gestures with 95.3% accuracy [3]. Weijie et al. also showed that the correctness rate of a gesture increases with the number of sensors [4]. Thus, increasing the number of sensors can improve the accuracy of hand motion estimation. However, increasing the number of sensors makes wearing a prosthetic hand more difficult and increases the cost of the prosthetic hand. For this reason, a method that enables accurate hand motion estimation with fewer sensors is needed.

Therefore, this paper proposes a method that enables accurate hand estimation from fewer sensors. For accurate hand motion estimation, this method uses super-resolution to increase the number of inputs without increasing the number of sensors. This method can be as effective as using many sensors, because super-resolution can generate more input signals from fewer inputs.

III. HAND MOTION ESTIMATION USING SUPER-RESOLUTION

A. Overview

Fig. 1 shows an overview of the proposed method. The inputs are multiple myoelectric signals. These signals are input to a deep neural network for super-resolution, and higher resolution signals are output. This neural network enables super-resolution by learning the relationship between signals at fewer points and those at more points. The super-resolved signals are then input to a deep neural network for hand motion estimation, and the estimation results are output. The estimated result is a class of hand motions.

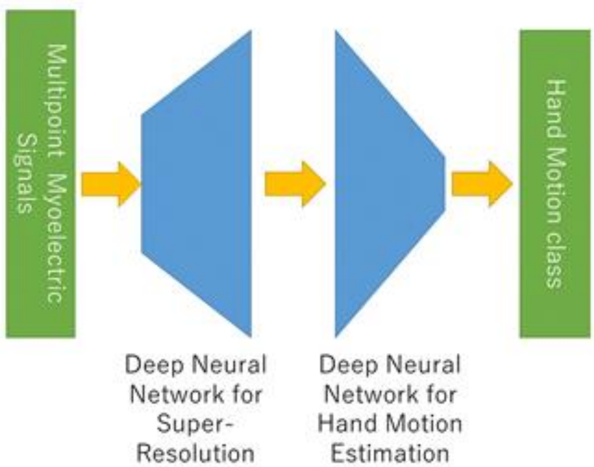

Fig. 1. Overview of the Proposed Method.

B. Super-Resolution for Multiple Myoelectric Signals

Super-resolution has been actively studied in the field of image processing [10]-[15]. Super-resolution in image processing generates a high-resolution image from a low-resolution image. In this study, this technique is used to obtain high-resolution myoelectric signals. Here, super-resolution uses the deep neural network shown in Fig. 2. Input is four-channel 512-sample EMG, output is 16-channel 512-sample EMG. The abbreviations in Fig. 2 mean the following.

- Conv: convolution layer
- PRelU: an activation function called parametric rectified linear unit [16].
- BN: batch normalization.
- Skip connection: a connection between two layers that passes information between them without transformation. Both skipped and un-skipped informations are available.
- Pixel Shuffler: It rearranges each pixel in the input feature map to output a high-resolution feature map [17].
- Residual block: a skip-connected block that learns a residual function by reference to the inputs of the layer, rather than learning a function that is unreferenced.

This neural network is based on SR ResNet [11]. SR ResNet keeps the height and width of the inputs the same size in the middle layer. In contrast, this method compresses the height by a quarter in the middle layer by convolution processing on the input. This is because both the height and width of the input are enlarged when using the Pixel Shuffler. By compressing the input data in advance so that the restored height is the same as the input data, it is possible to achieve more accurate super-resolution than with convolution or pooling processing. In Fig. 2, the output channel rows are replaced. This process is to make it easier to recover the original waveforms by super-resolution processing. The channel sequence is restored before input to the hand motion estimation network.

C. Hand Motion Estimation

Hand motion estimation uses deep neural networks. Fig. 3 shows a deep neural network for the hand motion classification. The input is a 16-channel 512-sample EMG, and the output is a label indicating the gesture. The abbreviations in Fig. 3 are the same as those used in Fig. 2. Max pool means a max pooling layer. Dropout is a method of preventing overfitting by inactivating a certain percentage of nodes while training a neural network [18]. The flatten layer converts the feature map into a one-dimensional vector. The softmax function transforms multiple output values into a sum of 1.0.

The proposed method assumes that hand motions are estimated every 0.25 seconds. A network of electromyograms with four input channels and 512 samples is also created for validation. Côté-Allard et al. [19] performed channel-by-channel convolution for hand motion estimation using multi-channel EMG. This method transforms 16 input channels with 512 sample inputs into 128 channels with a height of 64 and a width of 128 and applies convolution processing. This allows learning of numerical value changes over various times and channels by convolutional processing to achieve high accuracy.

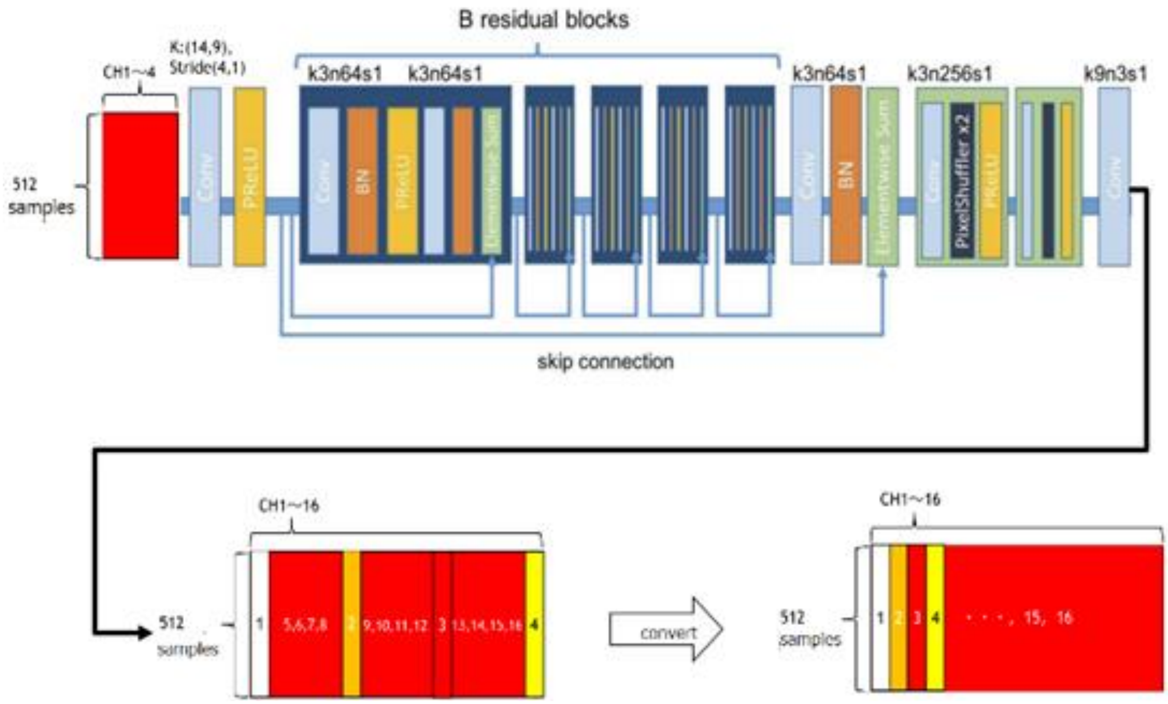

Fig. 2. Deep Neural Network for Super-Resolution.

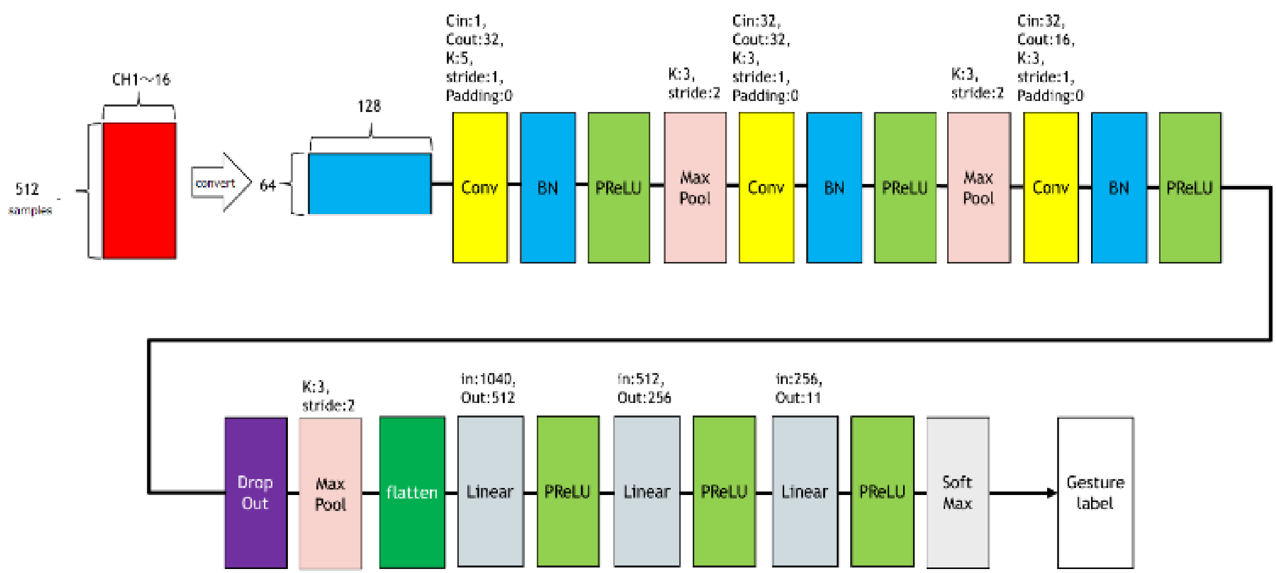

Fig. 3. Deep Neural Network for Hand Motion Estimation.

IV. Experiments

This section describes evaluation experiments using actual data to evaluate the effectiveness of the proposed method.

A. Dataset

The dataset in this study was acquired using the following devices. Fig. 4 shows the actual device used for this experiment. This device connects a myoelectric amplifier to a microcontroller board and sends its output to a PC. The electrode is a wet disposable ECG electrode (MSGLT-04). The electrodes were placed as shown in Fig. 5, taking into consideration the hand motions and the muscle arrangement. The reference electrode was placed at the elbow. The device was used to obtain myoelectric signals at 16 locations. This study uses the gestures in Fig. 6, as in the study by Côté-Allard et al. [19]. Each gesture is performed from a state of relaxation, and the device acquires EMG for 10 seconds during the gesture. The training data is the first nine seconds of data, and the remaining one second of data is used as the test data. To compensate for the small amount of data, the method of Lingfeng et al [20] was used. This produced 197131 training data and 16907 test data.

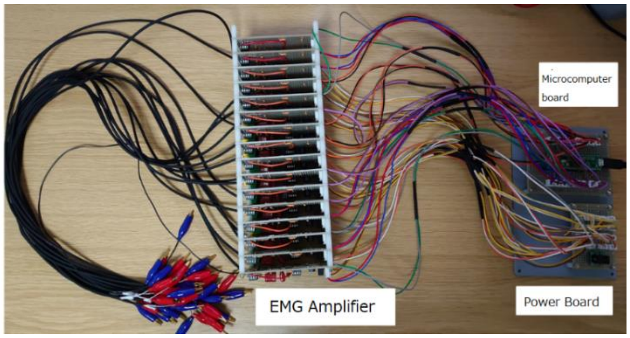

Fig. 4. EMG Measurement Device.

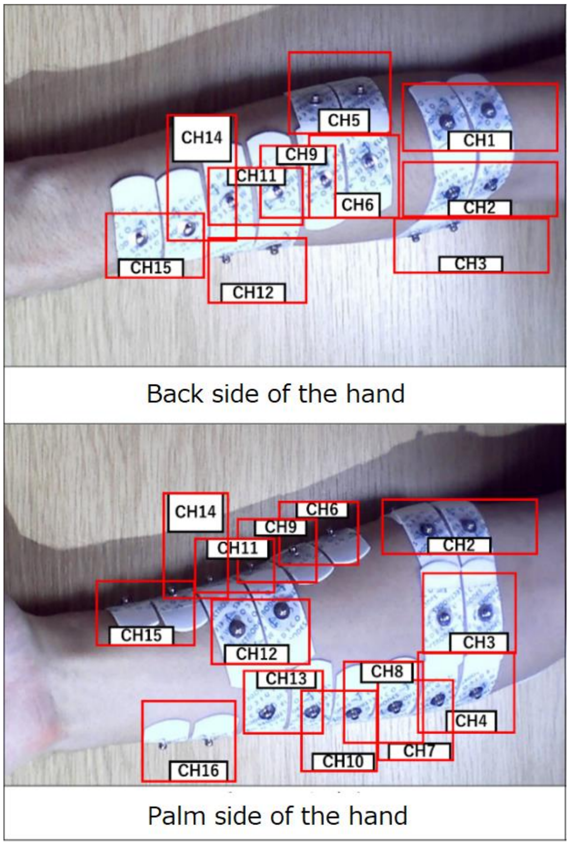

Fig. 5. Electrode Setting.

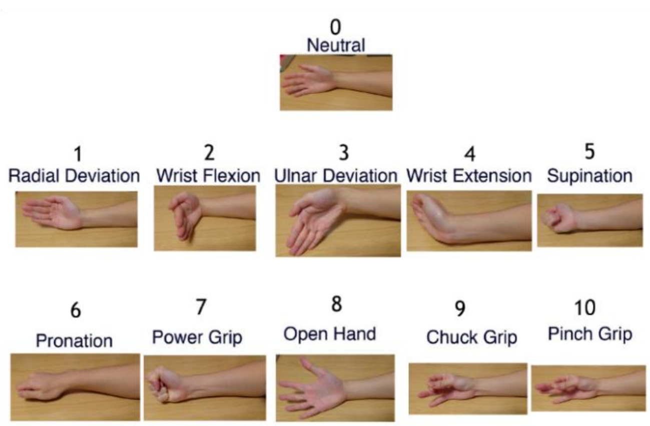

Fig. 6. Gesture Class.

B. Experiment of Super-Resolution

Here, the effect of super-resolution of myoelectric signals is verified using actual data. In this experiment, 4-channel 512-sample, EMG is the input, and the corresponding 16-channel 512-sample EMG is the supervised data. The mean squared error (MSE) was used as the loss function. The MSE is expressed in Equation (1).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y-t)^2 \tag{1}$$

N is the number of samples, and y and t are the output and supervised data, respectively. PSNR, which is also used in general image super-resolution, is used as a measure of the difference between the output and supervised data. PSNR is expressed by Equation (2). The smaller the difference between the output and supervised data, the larger the PSNR. Thus, the larger the PSNR, the better the output.

$$PSNR = 10log_{10}\left(\frac{1}{MSE}\right) \tag{2}$$

Using actual data, an effect of super-resolution of myoelectric signals was evaluated in an experiment. Experimental results of super-resolution learning on test data are shown in Fig. 7. Fig. 7 shows the variation of PSNR with the number of epochs. As with deep learning in general, PSNR increases as the epoch increases. This indicates that the accuracy of super-resolution improved without overfitting.

C. Experiment of Hand Motion Estimation

Here we examine the effectiveness of hand motion estimation using super-resolution. In this experiment, a 4-channel input, a 16-channel input, and a 4-channel input increased to 16 channels by super-resolution are compared. Fig. 8 compares the experimental results for each of these inputs. As noted in previous studies, 16 channels performed better than 4 channels on all evaluation items. This means that the accuracy of hand motion estimation is improved when the number of input EMG signals is large. The proposed method using super-resolution achieved better results than the 4-channel method in all evaluation items. This is because the proposed method increases the number of inputs from 4 channels to 16 channels. However, the proposed method was not able to reach 16 channels in all indicators. It is a reasonable result because the proposed method increases the number of channels to 16 with super-resolution, but it is only an estimated value.

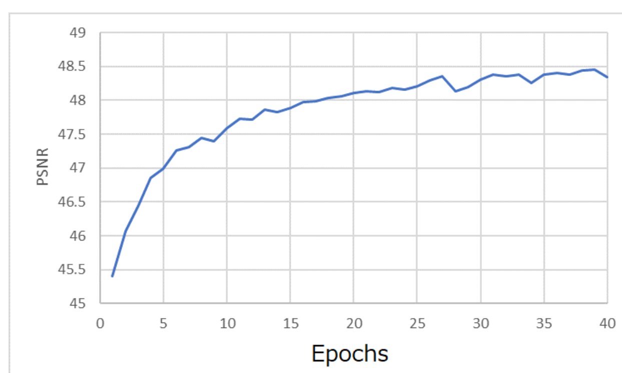

Fig. 7. PSNR for Number of Epochs.

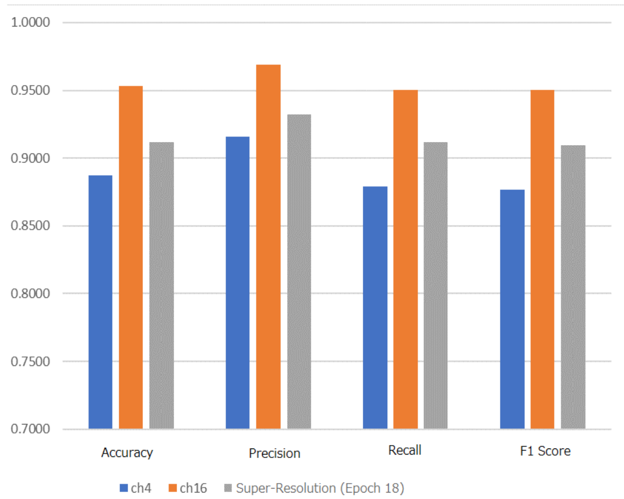

Fig. 8. Results of Hand Motion Estimation.

Based on the above experimental results, the proposed method is able to increase the accuracy of hand motion estimation even though the input is the same as the 4-channel method. This means that the knowledge acquired by the learning of super-resolution increased the accuracy of the estimation. Previous methods have increased accuracy by increasing the number of sensor inputs, however, the proposed method shows that accuracy can be increased without increasing the number of sensors.

Fig. 9 to 11 show the confusion matrices of the 4-channel, 16-channel, and proposed methods, respectively. Fig. 10 shows that the 4-channel input had more misclassifications than the other methods. For example, more of the correct label 8's were estimated as label 9, and more of the correct label 10's were estimated as label 9. The proposed method was able to classify many samples more accurately, although not as accurately as the 16-channel method. Labels 8, 9, and 10 are particularly misclassified, because these motions are close muscular motions.

The time required for hand motion estimation and super-resolution for 16907 EMGs, the test data, was 750 and 398 seconds, respectively. The time per EMG was (750 + 398) / 16907 ≈ 0.0680 seconds. Since the proposed method is designed to estimate hand motions every 0.25 seconds, the estimation speed is considered fast enough for practical use.

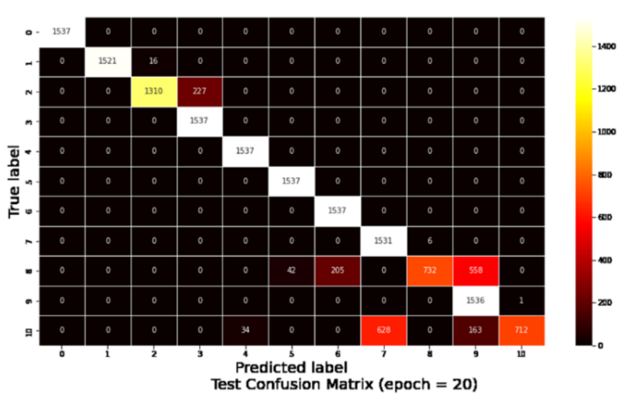

Fig. 9. Confusion Matrix of 4 Channel Input.

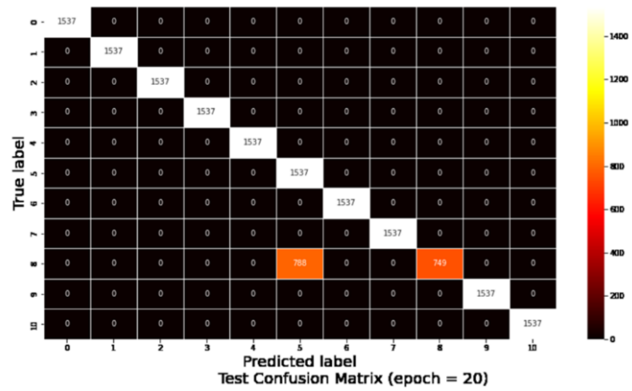

Fig. 10. Confusion Matrix of 16 Channel Input.

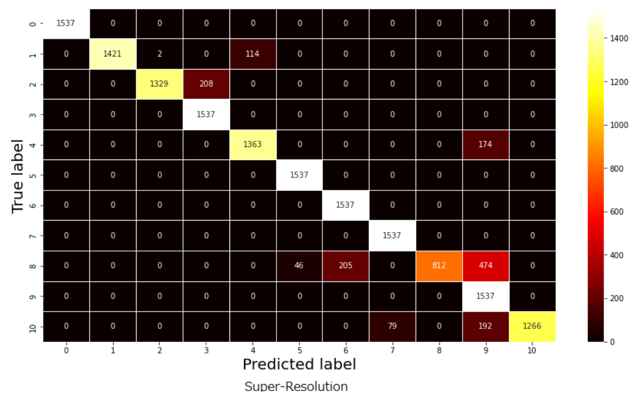

Fig. 11. Confusion Matrix of the Proposed Method.

V. Conclusion

In this paper, we proposed a method for estimating hand motions using super-resolution of multi-point surface electromyograms. The method uses a super-resolution network to obtain the relationship between EMGs with a small number of electrodes and those with a larger number of electrodes and estimates unknown high-resolution EMGs from new low-resolution EMGs. By inputting the estimated high-resolution EMG to the hand motion estimation network, hand motion estimation can be performed with fewer electrodes without loss of accuracy.

Hand motion estimation using super-resolution of multi-point surface EMG showed higher performance than hand motion estimation using four channels. This shows that super-resolution of electromyograms is available. This method may be used to reduce the number of electrodes in other EMG techniques as well. The accuracy was lower for gestures with the same moving muscles compared to other gestures. In this paper, the accuracy of super-resolution was evaluated using PSNR, which is used in super-resolution of general images. However, there was no relationship between the PSNR value and the results of hand motion estimation using electromyogram after super-resolution.

A future work is to reduce the number of samples per unit time of the input in super-resolution. Because this method focused on reducing the number of electrodes, the number of samples per unit time of the input was the same as that of the output. If the number of samples per unit time can be reduced, the devices used for EMG acquisition such as microcontroller and amplifier would become less expensive.

REFERENCES

[1] HUDGINS, Bernard; PARKER, Philip; SCOTT, Robert N., A new strategy for multifunction myoelectric control. IEEE transactions on biomedical engineering, 1993, 40.1: 82-94.

[2] Ali Raza Asif, et al., Performance evaluation of convolutional neural network for hand gesture recognition using EMG. Sensors, 2020, 20.6: 1642.

[3] Chen, J.; Bi, S.; Zhang, G.; Cao, G. High-Density Surface EMG-Based Gesture Recognition Using a 3D Convolutional Neural Network. Sensors 2020, 20, 1201.

[4] KE, Weijie, et al. Intersected EMG heatmaps and deep learning based gesture recognition. In: Proceedings of the 2020 12th International Conference on Machine Learning and Computing. 2020. p. 73-78.

[5] Lu, H.; Zou, Z.; Wu, X.; Shi, C.; Liu, Y.; Xiao, J. Biomimetic Prosthetic Hand Enabled by Liquid Crystal Elastomer Tendons. Micromachines 2021, 12, 736.

[6] Triwiyanto, T.; Caesarendra, W.; Purnomo, M.H.; Sułowicz, M.; Wisana, I.D.G.H.; Titisari, D.; Lamidi, L.; Rismayani, R. Embedded Machine Learning Using a Multi-Thread Algorithm on a Raspberry Pi Platform to Improve Prosthetic Hand Performance. Micromachines 2022, 13, 191.

[7] L. Vargas, H. Huang, Y. Zhu, D. Kamper and X. Hu, "Resembled Tactile Feedback for Object Recognition Using a Prosthetic Hand," in IEEE Robotics and Automation Letters, vol. 7, no. 4, Oct. 2022, pp. 10977-10984.

[8] Wang, S.; Zheng, J.; Zheng, B.; Jiang, X. Phase-Based Grasp Classification for Prosthetic Hand Control Using sEMG. Biosensors 2022, 12, 57.

[9] L. Vargas, H. Huang, Y. Zhu and X. Hu, Object Recognition via Evoked Sensory Feedback during Control of a Prosthetic Hand, in IEEE Robotics and Automation Letters, vol. 7, no. 1, Jan. 2022, pp. 207-214.

[10] DONG, Chao; LOY, Chen Change; TANG, Xiaoou. Accelerating the super-resolution convolutional neural network. In: European conference on computer vision. Springer, Cham, 2016. p. 391-407.

[11] LEDIG, Christian, et al. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. p. 4681-4690.

[12] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet and M. Norouzi, "Image Super-Resolution Via Iterative Refinement," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022doi: 10.1109/TPAMI.2022.3204461.

[13] Junjun Jiang, Chenyang Wang, Xianming Liu, and Jiayi Ma. 2021. Deep Learning-based Face Super-resolution: A Survey. ACM Comput. Surv. 55, 1, Article 13 (January 2023), 36 pages.

[14] Honggang Chen, Xiaohai He, Linbo Qing, Yuanyuan Wu, Chao Ren, Ray E. Sheriff, Ce Zhu, Real-world single image super-resolution: A brief review, Information Fusion,Volume 79, 2022, Pages 124-145.

[15] Yiqun Mei, Yuchen Fan, Yuqian Zhou; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 3517-3526.

[16] HE, Kaiming, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision. 2015. p. 1026-1034.

[17] SHI, Wenzhe, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 1874-1883.

[18] SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research. Journal of Machine Learning Research, 1929, 1929-1958.

[19] Côté-Allard, Ulysse, et al. "Interpreting deep learning features for myoelectric control: A comparison with handcrafted features." Frontiers in bioengineering and biotechnology 8 (2020): 158.

[20] Xu, L., Chen, X., Cao, S., Zhang, X., Chen, X., Feasibility study of advanced neural networks applied to sEMG-based force estimation. Sensors, 2018, 18(10), 3226.