International Journal of Control, Automation and Systems 19(3) (2021) 1443-1450
http://dx.doi.org/10.1007/s12555-019-0802-1
ISSN:1598-6446 eISSN:2005-4092
http://www.springer.com/12555

Classification of Hand Gestures Based on Multi-channel EMG by Scale Average Wavelet Transform and Convolutional Neural Network

Do-Chang Oh* and **Yong-Un Jo**

Abstract: Predicting and accurately classifying intentions for human hand gestures can be used not only for active prosthetic hands, rehabilitation robots, and entertainment robots but also for artificial intelligence robots in general. In this paper, first of all, source data of three hand gestures of grasping and three hand gestures of sign language are acquired by using the armband combined with eight sEMG (surface Electromyography) sensors. To classify these hand gestures, basically simple CNN (convolutional neural network) models with raw data, short-time Fourier transform (STFT), wavelet transform (WT), and scale average wavelet transform (SAWT) are applied, and their performances are compared. Finally, it is shown that by using a CNN with SAWT images, the accuracy can be improved up to 94.6% for selected hand gestures with higher accuracy and lower computational burden than conventional multi-channel STFT or WT.

Keywords: Accuracy, classification, CNN, hand gestures, scale average wavelet transform (SAWT), sEMG.

1. INTRODUCTION

Robotics and artificial intelligence can be leveraged to increase the autonomy of people living with disabilities. This is accomplished, in part, by enabling users to seamlessly interact with robots to complete their daily tasks with increased independence [1].

Electromyography (EMG)-based techniques have been used to recognize hand gestures and estimate finger motion for various things such as non-verbal communication with sign language and gesture based interfaces. They are also very useful in assisting with system control [2]. EMG signal is a biosignal with several potential applications, including diagnosis of neuromuscular diseases, control of systems such as prosthetics, robots, entertainment devices. The use of the EMG signal with efficient recognition of hand gestures can help develop a direct machine interface, making it possible to increase the autonomy of people with special needs [3–5].

The relationship between EMG signals and body movements has not been established completely, though it is known that the signals have a high correlation to body movements. For data acquisition of EMG signals related to upper limb movements, a surface EMG (sEMG) sensor is typically used on a person's forearm. In particular, Myo-armband (Thalmic Labs Inc.) is a commercially available sensor device that has often been used to predict human hand gestures [1–4, 6, 7]. Many studies are mainly aimed at rehabilitation of paralyzed patients and prosthetic assistance of amputation patients. Therefore, it is necessary to make the feature extraction and classification of the user's EMG signals obtained through the Myo-armband.

In recent years, studies on increasing accuracy and real-time performance for hand gesture prediction and control have been actively conducted with the introduction of machine learning and deep learning [1–11]. In [3], Fonseca *et al*. aimed at recognizing the hand gestures pattern motivated by human-robot interaction. They used the artificial neural network (ANN) with pre-processing for the feature extraction. Asai *et al*. [2] showed the results of classifying four hand gestures using the convolutional neural network (CNN) with wavelet transform input data, including an automatic labeling system. In [4], Na *et al*. showed the CNN (AlexNet) based on spectrograms has a higher classification accuracy than the support vector machine (SVM). Batzianoulis *et al*. proposed an EMG-based learning approach that decodes the grasping intention at an early stage of reach-to-grasp motion [5]. Atzori *et al*. tested their proposed deep learning methods for natural control of robotic hands via sEMG using a large number of intact subjects and amputees [8].

On the other hand, deep learning with a recurrent CNN model was used for EMG-based estimation of limb movement [9]. The end-to-end deep learning derived from the time-frequency representation of EMG signals was proposed for identifying normal and aggressive actions [10].

Manuscript received September 23, 2019; revised February 24, 2020; accepted June 29, 2020. Recommended by Associate Editor Jong-Han Kim under the direction of Editor Doo Yong Lee. This paper was supported by the Konyang University Research Fund in 2019.

Do-Chang Oh and Yong-Un Jo are with the School of Biomedical Engineering, Konyang University, 158 Gwanjeodong-ro, Seo-gu, Daejeon 35365, Korea (e-mails: {docoh, 19806504}@konyang.ac.kr).
* Corresponding author.

Although the structure is complicated, the transfer learning augmented CNN scheme has been shown to enhance three networks' performance on the two datasets [1]. In the previous work [11], a CNN model with two convolutional layers and two fully connected layers was directly constructed through function and parameter settings. Five hand gestures were predicted using the sEMG input data and the CNN model.

Deep learning is a branch of the artificial neural network. It has a unique hierarchical structure and the ability to extract high-level features. Deep learning networks have been used widely in several applications, such as medical diagnoses, biomedical signal classifications, and speech and fault diagnosis [12–15]. On the other hand, to extract the feature, frequency domain signals using the Fourier transform and time-domain signals were used. Since deep learning deals efficiently with images, scientists have resorted to transforming signals into visual representations based on time-frequency representation.

Time-frequency can disclose characteristic signal patterns. It is also a powerful tool for characterizing medical signals [1, 2, 4, 9, 10, 12–18]. Two linear time-frequency transformations are utilized on EMG signals as image inputs to the CNN architecture: the short-time Fourier transform (STFT) spectrogram and the wavelet transform (WT) scalogram. These types of images are another form of raw signal feature representation [10].

The spectrogram images were used to train a CNN for some medical identifying of diagnoses such as automatic atrial fibrillation detection, arrhythmia, sleep disorder, motor impairment neural disorder, and clinical brain death using time-frequency images of biosignal (ECG, EEG, and EMG) [4, 9, 10, 13–15]. In addition, research has been conducted to classify various human behaviors using wavelet transformed scalogram image data and deep learning [1, 2, 10, 16–18].

In this paper, source data of three hand gestures of grasping and three hand gestures of sign language are acquired using the armband combined with eight sEMG sensors. To classify these hand gestures, basic simple CNN models with raw data, STFT, WT, and scale average wavelet transform (SAWT) are applied, and their performances are compared. As a result, it is shown that by using the newly proposed SAWT in this paper, the accuracy can be improved up to 94.6% for selected hand gestures with a lower computational burden than conventional multi-channel STFT [4] or WT [2].

2. MAIN RESULTS

2.1. Hand gesture data processing

To receive signals from the sEMG sensor and classify the shape of hand gestures, we selected the hand gestures that are most commonly formed to express human intent or hold objects. Hand gesture selection can be done in a variety of experimental ways, but here, as shown in Fig. 1, six representative hand gestures to grasp objects or to express human intention have been chosen as "fist", "pinch", "card", "ok", "victory", and "good". There can be many issues depending on the area of use, including how many different hand gestures are classified, or how precisely classified, or how quickly classified. In the case of the active prosthetic hand, which can reflect the person's intention, it can be expected to play some necessary roles in the three hand gestures of grasping and the three hand gestures of expression.

(a) Hand gestures of grasping.

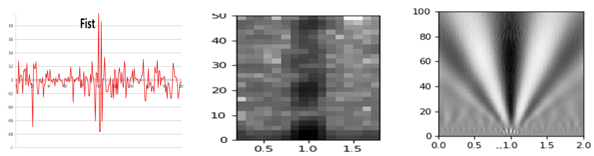
(b) Hand gestures of expression.

Fig. 1. Selected hand gestures for grasping or intention expression.

EMG signals show some variability depending on the subject and environment, but most of the rehabilitation control devices using intention signals from EMG signals, such as prosthetic hands and wearable robots, are custom-made and need to be adjusted for one user. Here, by learning and testing from the three subjects' data, classification accuracy is improved and analyzed using the proposed CNN classifier. Therefore, when receiving data using the Myo armband in Fig. 1, the subject starts with their fingers spread and performs the six selected gestures. Also, it is performed 200 times per one gesture, 1200 times in total. When one operation is taken once, 50-200 samples are taken at intervals of 2 seconds and used as one dataset of CNN (sampling frequency 5 Hz-125 Hz).

2.2. Scale average wavelet transform and CNN

The deep learning network has been considered in a number of applications as a form of end-to-end learning wherein feature extraction, pre-processing, and classification are conducted directly. There are various classification methods such as SVM (support vector machine), hidden Markov models, and the neural network for hand motion recognition and estimation. Since feature extraction generally dominates classification performance, identifying appropriate features is essential for successful classification. Fig. 2 compares the source image (raw data) with the image of STFT (short-time Fourier transform) and WT (wavelet transform) for the fist gesture in Fig. 2. Here, the horizontal axis is time (sample sequence), and the vertical



(a) Source signal. (b) STFT image. (c) WT image.

Fig. 2. Comparison the source with STFT and WT.

axis is EMG amplitude (a), frequency (b), and scale (c), respectively.

Time-domain features such as EMG, EMG integration, mean, and variance have been used to capture stationary conditions of EMG activity, and time-frequency domain features by STFT and wavelet transform have been used to capture transient conditions. We use CNN as a classifier to estimate hand gestures from EMG signals. CNN is a kind of multi-layer neural network that shows successful results in image classification, classification for EMG signal is treated as a 2D image recognition problem. The 2D images obtained by the wavelet transform of EMG signals are input data of CNN. The structure of CNN is based on 2-hidden layers and 2-fully connected layers, as shown in Fig. 3. To improve system performance, various techniques such as Relu activation function, Xavier initialization function, and dropout were applied. In Fig. 3, CNN inputs are shown as EMG raw image data and wavelet transform image data, and they are explained in detail later.

sEMG signals were processed using continuous-time wavelet transform (CTWT). CTWT is a time-frequency analysis method that quantifies temporal changes in the frequency content of non-stationary signals without losing resolution in time or frequency [17]. CTWT of the input signal $x(t)$ is defined as the inner product,

$$W(a,b) = \int_{-\infty}^{\infty} x(t)\psi^{*}_{a,b}(t)dt, \quad a \neq 0, \tag{1}$$

where the basis function $\psi_{a,b}(t)$ is the mother wavelet, featured by scale and translation parameters (time shifting), a and b, respectively.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right). \tag{2}$$

$\psi_{a,b}(t)$ is obtained by the mother wavelet function $\psi(t)$ at time b and scale a. In terms of frequency, the multiresolution analysis provides the global information of the signal corresponding with low frequencies and the detailed information associated with high frequencies.

In discrete-time wavelet transform (DTWT), a signal is analyzed with a small number of scales with varying translations at each scale. A critical sampling of the CTWT $W(a,b)$ is obtained by substituting a by j and b by k in (2), where j and k are integers representing the scale and translation. Upon this substitution, $\psi_{j,k}(t)$ is a discrete set

$$\psi_{j,k}(t) = \frac{1}{\sqrt{j}}\psi(\frac{t-k}{j}) \tag{3}$$

of child wavelets for a given mother wavelet $\psi(t)$. The term critical sampling is used to ensure that a minimum number of wavelet coefficients are retained to represent all the information present in the original function. The integral with t can be replaced with the formula of discrete sum; thus, DTWT is

$$W(j,k) = \sum_{n=-\infty}^{\infty} x(n)\frac{1}{\sqrt{j}}\psi(\frac{n-k}{j}). \tag{4}$$

In CTWT, transform coefficients are found for every (a,b) combination, whereas in DTWT, transform coefficients are found only at very few points.

Fig. 4 shows wavelet scalograms according to the change of scale parameter for 8-channel data of the fist hand gesture.

The horizontal axis represents 200 samples for 2 seconds as a shift parameter, and the vertical axis represents the scale change as a frequency variation. The smaller the scale value, the higher the time resolution and the lower

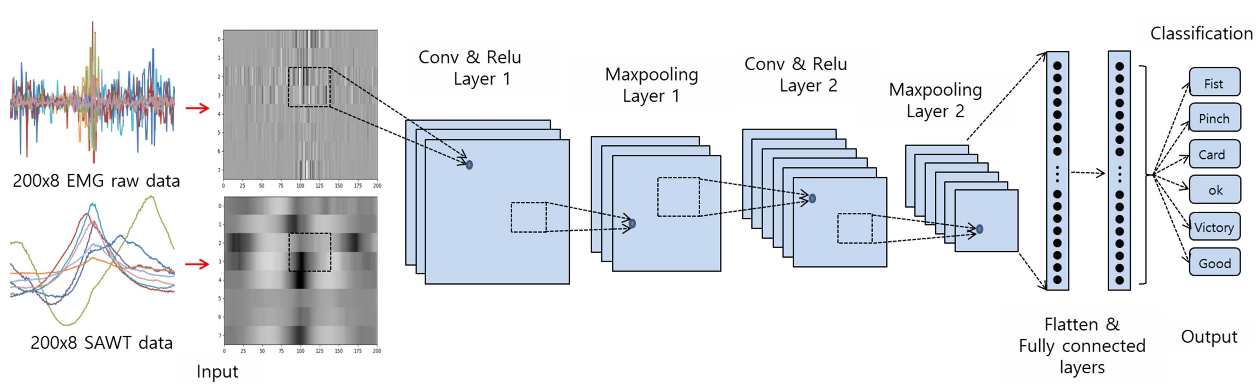


Fig. 3. The structure of CNN with EMG image data.

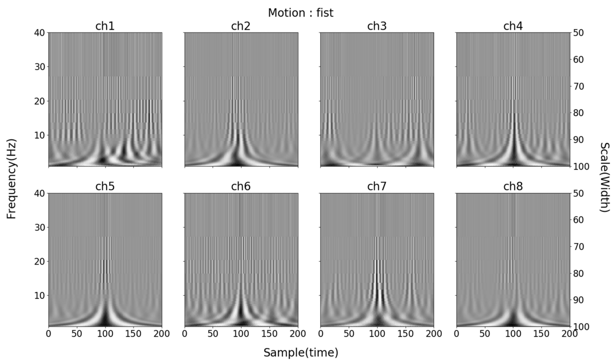


Fig. 4. Wavelet scalogram for fist hand gesture.

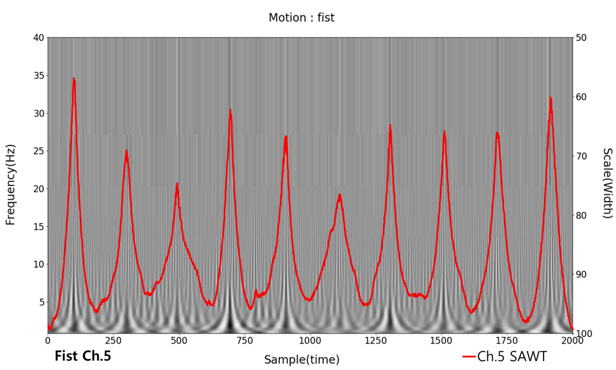


Fig. 5. SAWT on scalogram (red line graph).

the frequency resolution, and vice versa. In addition, we have further enhanced the feature by averaging the scalogram values according to scale value change to make the features of the image more prominent.

If the number of scale and input sequences is J and N, the scale average wavelet transform (SAWT) is defined as follows:

$$W(k)_{SAWT} \equiv \frac{1}{J}\sum_{j=1}^{J}\sum_{n=0}^{N-1} x(n)\frac{1}{\sqrt{j}}\psi(\frac{n-k}{j}), \tag{5}$$

where k ($k = 0, 1, 2, ..., N-1$) is the time shifting parameter, j is the j-th scale parameter ($j = 1, 2, ..., J$) and n ($n = 0, 1, 2, ..., N-1$) is the sequence number of mother function and input, and these parameters are all integers.

In Fig. 5, the average value (red line graph) of scalogram data from $j = 1$ to $j = 200$ ($J = 200$) is plotted on the scalogram for the fist gesture. Hereafter, this averaged method is referred to as the Scale Average Wavelet Transform (SAWT). In Fig. 6, it shows the formation of one frame image consisting of 8-channel image data generated by SAWT.

Fig. 7 shows graphs of scalograms with scale values $f/2$ and $2f$, and SAWT up to $2f$, respectively, while repeating the fist operation twice for 4 seconds, where f is the sampling frequency (100 Hz). Also, the reason for choosing two channels of ch. 5 and ch. 6 is because it is

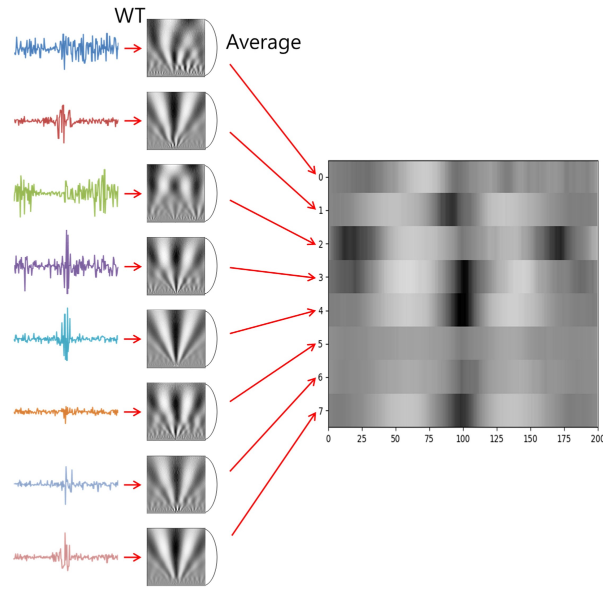


Fig. 6. 8-channel image data generated by SAWT.

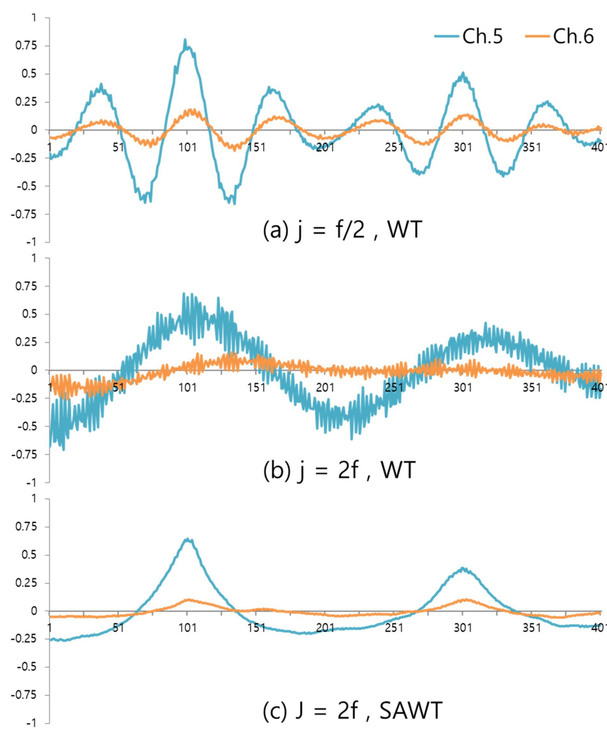


Fig. 7. Wavelet transform with scale $f/2$ and $2f$, and SAWT up to $2f$.

possible to compare which channel is more active for a certain gesture and classify gestures from the result of the feature.

When the scale is relatively small $f/2$, the first graph exhibits detailed motion, whereas the overall trend is difficult to see. In the case of $2f$ with a large scale, the overall trend is easy to see but includes detailed noise. Therefore,

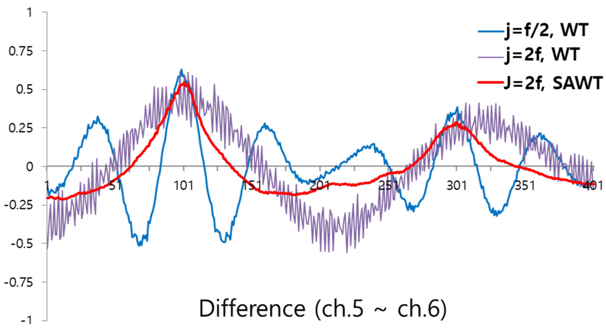


Fig. 8. Difference between ch.5 and ch.6 in WT and SAWT.

in the last graph, WT data is averaged over the scale axis so that the trend and detail are properly included.

Also, as shown in Fig. 8, when the average value of a scalogram is used, it is easy to distinguish between channel 5 which is relatively active and channel 6, which is not activated from the difference of the SAWT between two channels.

3. EXPERIMENTAL RESULTS

3.1. Sampling frequency and scale bound of SAWT

In the experiment, SAWT was performed using the "Morlet" mother function, the upper bounds of scale value are selected as $f/2$, f, $2f$, and $4f$, where f is the EMG sampling frequency from 25 Hz to 125 Hz. Since human hand gesture signal has a very low frequency band, the sampling frequency of EMG signal reaches up to 150 Hz or less. Here, sampled the EMG signals were obtained below 125 Hz considering the band frequency and overfitting of CNN. Therefore, the scale value can be minimum (2.5, 5, 10, 20) to maximum (62.5, 125, 250, 500). The x-axis is the sampling frequency and the y-axis represents deep learning test accuracy (%). In Fig. 9, Original is EMG raw data without WT, and $J = f/2$, f, $2f$, and $4f$ are accuracy data obtained by applying SAWT to each upper bound of scale value.

At all sampling frequencies in Fig. 9, accuracy was the lowest for scale $J = f/2$ and the highest for $J = 2f$. Also, Since sampling frequency is high, and the number of data increases, test accuracy tends to increase to a certain level and fall again. The highest accuracy of the original is 86% at 25 Hz, 81.3% at 50Hz for $J = f/2$, 94% at 100 Hz for $J = 2f$, and 93.4% at 100 Hz for $J=4f$. As a result, for all frequencies, SAWT with $J = 2f$ shows higher test accuracy than other scale value bound cases. When the scale value bound is $2f$, and the sampling frequency is 100Hz, the highest test accuracy of 94% is obtained.

The confusion matrix showing the accuracy of the output class estimated for the target class is shown in Fig. 10. Note that if the target is "victory", it may be wrongly recognized as an image of "ok" or "card". Especially, the "victory" was frequently seen as "ok".

3.2. CNN performances with different input images

Data were collected from one female and two males under the same conditions (electrode position, arm position, posture). All the subjects performed 200 times per one gesture at 100Hz sampling rate, and STFT, WT, SAWT ($J = 200$) are obtained by using raw data received from 8ch Myo-armband, and four types of input images are composed in a time domain and time-frequency domain. Similar to the methods of [2] and [4], the raw data, STFT, and WT data are each made of one corresponding input image, completed in the form of a 2×4 block matrix using eight images of each channel. In the case of SAWT, input image size is greatly reduced because the scale average is calculated for each channel [2, 4]. Of course, this is an advantage in reducing the computational burden of the system, and as the number of channels decreases, the size of input data can be further reduced.

Four types of input images for deep learning are obtained from three subjects. Using the proposed CNN model, 83% (200 sets) of each subject's data is trained and the remaining 17% (40 sets) is tested. Table 1 shows the test accuracy(%) when CNN models with four dif-

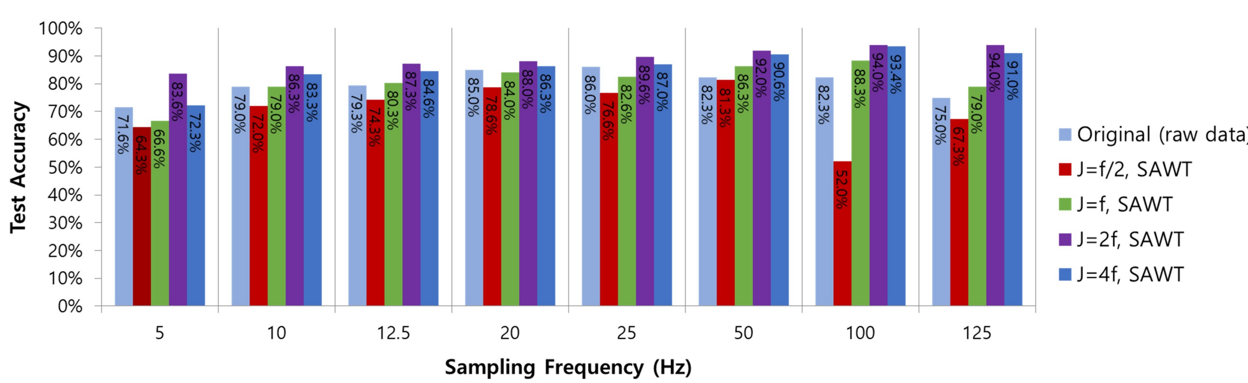


Fig. 9. CNN test accuracy at all sampling frequencies.

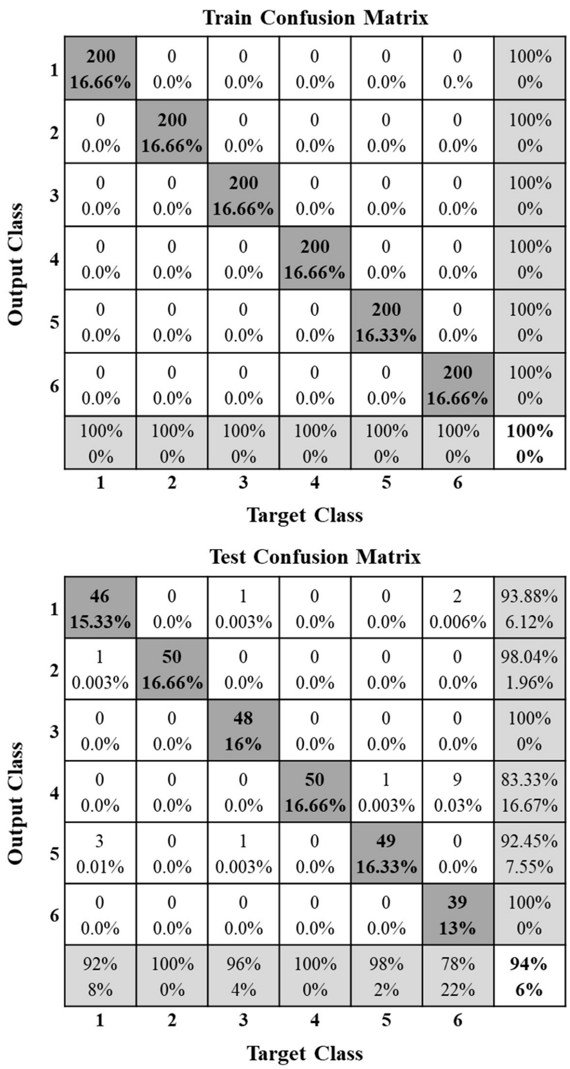


Fig. 10. SAWT up to 2f confusion matrix.

ferent input images were applied to three subjects (A, B, and C). STFT includes cases with four different window lengths and overlap lengths (window length–overlap length in Table 1 and Fig. 11, and WT includes three different widths (scales) range. Fig. 11 graphically shows the test average accuracy. As a result, when compared with raw data (89.44%), STFT (92.5%), and WT (92.1%), the CNN model with SAWT ($J = 200$) input image showed the highest average accuracy of 93.89% (the maximum accuracy is 94.6%).

EMG input image to CNN mainly uses STFT and WT in the time-frequency domain rather than using raw data in the time domain. In the case of using STFT, large input data and complex network model are used, but the accuracy result for ten kinds of gesture classification is good at 94% [4]. Literature [2] shows the results of classifying four hand gestures with 83% accuracy using WT data, including an automatic labeling system. On the other hand,

Table 1. Test accuracy(%) for three subjects.

Subject	Raw	STFT 15-10	STFT 28-20	STFT 50-40	STFT 100-75	WT 4	WT 8	WT 16	SAWT 200
A	90.0	89.2	91.7	91.7	90.4	86.7	88.7	87.1	94.6
B	87.1	94.2	93.3	93.3	88.3	93.7	94.2	94.2	94.2
C	91.2	92.9	92.1	92.5	90.8	90.4	93.3	93.3	92.9
Average	89.4	92.1	92.4	92.5	89.9	90.3	92.1	91.5	93.9

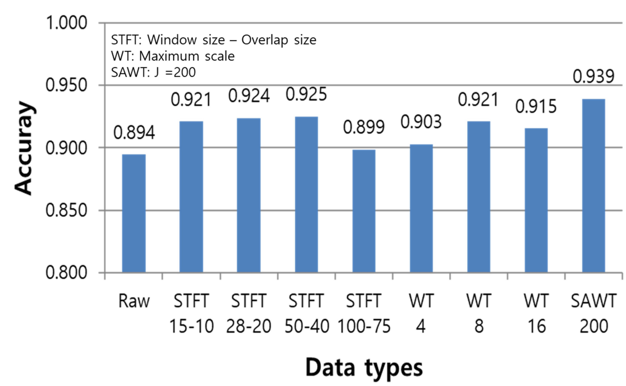


Fig. 11. Test average accuracy for CNN models with four input images (raw data, STFT, WT, SAWT).

STFT and WT images generated from EMG were applied to CNN to distinguish between normal and aggressive action. The maximum accuracy was 94.6% (EMG sensor is different from Myo-armband) [10].

This paper focuses on showing the superiority of SAWT. Notably, the CNN approach using the SAWT presented in this paper has higher accuracy with less data size than the three cases of raw data, STFT data, and conventional WT under the same experimental condi-tions. However, as in many pieces of literature, experimental conditions may vary, such as the number and shape of gestures, the type and size of input data, the number of subjects, and SEMG sensors. Therefore, it is difficult to compare the results with previous studies directly.

In Fig. 12, all data from the three subjects are learned and the test accuracy of each subject is shown for the CNN with SAWT ($J = 200$) image.

In single-user applications, such as personal prostheses or wearable robots, learning from many subjects' data may yield higher or lower test accuracy than the individual subjects, even though it depends on the subject. However, when the proposed method is applied to a multi-user robot, the overall test accuracy should be high, and the deviation of the individual accuracy be small. In future research, it is necessary to experiment, analyze, and improve with more subjects to expand the application of the CNN algorithm or increase verification reliability.

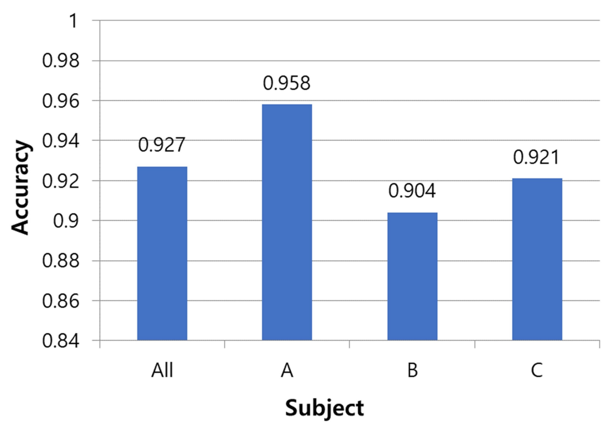


Fig. 12. CNN test accuracy for all subjects training and individual testing with SAWT image.

4. CONCLUSION

Source data of three hand gestures of grasping and three hand gestures of expression were acquired by using the armband combined with eight sEMG sensors. To classify these hand gestures, raw data, WT, and SAWT (scale average wavelet transform) images were applied, and a basic deep learning classifier CNN was used. In conclusion, it was demonstrated that by using a CNN with the newly proposed SAWT, the accuracy could be improved to 94.6% for selected hand gestures with higher accuracy and lower computational burden than conventional multi-channel STFT or WT.

Human EMG data collected in real-time according to the user's environment is not stable. Therefore, in the future, we will study hand gestures classification based on real-time collected data and SAWT, and also define human work schedules and study classification of hand movements based on them. Furthermore, we need to study whether deep learning classifiers can learn more people's data to find and classify features in specific hand gestures from each EMG signal.

REFERENCES

[1] U. C. Allard, C. L. Fall, A. Drouin, A. C. Lecours, C. Gosselin, K. Glette, F. Laviolette, and B. Gosselin, "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 760-771, 2019.

[2] K. Asai and N. Takase, "Finger motion estimation based on frequency conversion of EMG signals and image recognition using convolutional neural network," *Proc. of 17th International Conference on Control, Automation and Systems (ICCAS)*, Oct. 18-21, Jeju, Korea, 2017.

[3] M. G. B. Fonseca, A. G. S. Conceicao, and E. F. S. Filho, "Artificial neural networks applied to the classification of hand gestures using electromyo- graphic signals," *Proc. of Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics*, Curitiba, Brazil, IEEE 2017.

[4] D. Na, L.-Z. Liu, X.-J. Yu, Q. Li, and S.-C. Yeh, "Classification of multichannel surface-electromyography signals based on convolutional neural networks," *Journal of Industrial Information Integration*, vol. 15, pp. 201.206, Sept. 2019.

[5] I. Batzianoulis, S. El-Khoury, E. Pirondini, M. Coscia, S. Micera, and A. Billard, "EMG-based decording of grasp gestures in reaching-to-grasping motions," *Robotics and Automation Systems,* vol. 91, pp. 59-70, Jan. 2017.

[6] J. Luo, C. Liu, and C. Yang, "Estimation of EMG-based force using a neural-network-based approach," *IEEE Access*, vol. 7, pp. 64856-64865, 2019.

[7] A. Boyali and N. Hashimoto, "Spectral collaborative representation based classification for hand gestures recognition on electromyography signals," *Biomedical Signal Processing and Control*, vol. 24, pp. 11-18, 2016.

[8] M. Atzori, M. Cognolato, and H. Muller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers in Neurorobotics*, vol. 10, no. 9, Sept. 2016.

[9] P. Xia, J. Hu, and Y. Peng, "EMG-based estimation of limb movement using deep learning with recurrent convolutional neural networks," *Artificial Organs,* vol. 42. no. 5, pp. 67-77, Oct. 2017.

[10] H. Alaskar, "Deep learning of EMG time-frequency representations for identifying normal and aggressive actions," *International Journal of Computer Science and Network Security*, vol. 18, no. 12, Dec. 2018.

[11] Y. U. Jo and D. C. Oh, "Study on the EMG-based grasp gesture classification using deep learning and application to active prosthetics," *Journal of Institute of Control, Robotics and Systems* (in Korean), vol. 25, no. 3, pp. 229-234, 2019.

[12] G. Ruffini, D. Lbanez, M. Castellano, L. D. Vall, A. S. Frisch, R. Postuma, J. F. Gagnon, and J. Montplaisir, "Deep learning with EEG spectrograms in rapid eye movement behavior disorder," *Frontiers in Neurology*, vol. 10, Article 806, July, 2019.

[13] S. Chauhan and L. Vig, "Anomaly detection in ECG time signals via deep long short-term memory networks," *Proc. of IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 1-7, 2015.

[14] L. Yuan and J. Cao, "Patients' EEG data analysis via spectrogram image with a convolution neural network," *Proc. of International Conference on Intelligent Decision Technologies*, pp. 13-21, 2017.

[15] A. Vilamala, K. H. Madsen, and L. K. Hansen, "Deep convolutional neural networks for interpretable analysis of EEG sleep stage scoring," *Proc. of IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2017.

[16] H. Chen, P. Zhang, H. Bai, Q. Yuan, and Y. Yan, "Deep convolutional neural network with scalogram for audio scene modeling," *Interspeech 2018*, pp. 3304-3308, Hyderabad, India, 2018.

[17] J. Kilby and K. Prasad, "Continuous wavelet analysis and classification of surface electromyography signals," *International Journal of Computer and Electrical Engineering,* vol. 5, no. 1, February 2013

[18] A. Strazza, F. Verdini, L. Burattini, S. Fioretti, and F. D. Nardo, "Time-frequency analysis of surface EMG signals for maximum energy localization during walking," *EMBEC & NBC 2017, IFMBE Proceedings*, vol. 65, 2017.

Do-Chang Oh received his B.S., M.S., and Ph.D. degrees in electronics from Kyungpook National University in 1991, 1993, and 1997, respectively. He was with the University of Florida as a Courtesy Associate Professor for one year from July 2007. He is currently a professor at the School of Biomedical Engineering, Konyang University. His research interests include robust control, model reduction, rehabilitation robot, and biomedical applications of deep learning.

Yong-Un Jo received his B.S. degree in biomedical engineering from Konyang University in 2019. He is currently studying for a master's degree at Konyang University. His research interests include ANN(artificial neural network), rehabilitation robot, and biomedical applications of deep learning.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.