

Feature Extraction of Facial Electromyography (EMG) Signal for Aceh Languages Speech using Discrete Wavelet Transform (DWT)

Darma Setiawan Putra^{a,1,*}, Yuril Uumbu Woza Weru^{b,2}

^{a,b} Informatics Engineering, Politeknik Aceh Selatan, Merdeka Street, Tapaktuan, 23711, Indonesia

¹darma.poltas@gmail.com*; ²yurilweru@gmail.com

* corresponding author

ARTICLE INFO

ABSTRACT

Article history:
Accepted

Keywords:

Facial electromyography
FEMG
Speech recognition
Discrete wavelet transform
Feature extraction

The facial electromyography (FEMG) signal is a signal that occurs in the muscles of the contracted human face. This FEMG signal is one of the techniques used to study human speech recognition. It can be acquired by placing an electrode surface on the skin around the facial articulation muscle. Three types of muscles in this study are the masseter, rigorous and depressor muscle. This study aims to extract and analyze the features in the FEMG signal. The extraction method is the discrete wavelet transform (DWT). The type of wavelet transform is Daubechies2 with level 5. After extraction and analysis of FEMG signals, the FEMG signal pattern for each spoken word indicated by differences in the approximation and detail coefficient of the FEMG signal. In addition, the level of difference in the FEMG signal pattern is also indicated by the histogram of the approximation coefficient of the FEMG signal. Thus, the discrete wavelet transform method can be used as one of the methods for extracting the FEMG signal feature in a human facial electromyography (FEMG) signal.

Copyright © 2019 Politeknik Aceh Selatan.
All rights reserved.

I. Introduction

The development of science and technology has significantly influenced in signal processing technology. The human has electromyography (EMG) signal in the arm, leg, and facial muscle. The EMG signal in the face is called the facial electromyography (FEMG) signal. The FEMG is usually used to measure the level of emotional expression in a human face. This FEMG signal is recorded by attaching a surface electrode to the skin of the face to find out the emotional expression. The emotional expressions that are usually measured are expression of pleasure, sadness, and anger. The facial muscles used to measure the EMG signals are corrugator supercili muscle, zygomaticus major and orbicularis oculi [1][2][3]. This emotional expression is interpreted by measuring the magnitude of facial EMG signals. Emotional expression can also be associated with specific FEMG signal pattern. The FEMG signal can also show physiological conditions as an objective measure of facial muscle activity. When there is a change in facial expression, it will cause contractions in the facial muscle. This FEMG signal can be used to recognize human facial gesture for assistive technology and rehabilitation [4][5][6]. In addition, measuring emotional expression and facial gesture, this FEMG signal can measure signal pattern when the human speech in daily communication [7][8][9][10]. When human speaks the languages, the articulation muscles around the mouth contracts that can cause FEMG signal. This study aims to analyze the feature of the FEMG signal when speaking in Aceh language. So that this feature can be used for pattern recognition of FEMG signal in human facial muscle.



II. Related Studies

Some previous studies about feature extraction methods based on facial electromyography (FEMG) signals have been carried out. The feature extraction method using a discrete wavelet transform for the application not only signal extraction but also image extraction [11]. Previous research by [12] has conducted a study to extract the EMG signal features for Spanish syllable pronunciation using the Fast Fourier Transform (FFT) method. The resulted features are 20 components namely root mean square, average amplitude, maximum amplitude, kurtosis, mel-frequency cepstral coefficient as many as 13 values, mean absolute value, zero crossing. The syllables used are 5 units, namely vowel, labials, dentals, palatals, velars, and alveolar. The next research conducted by [8] to extract the EMG signals for Thai pronunciation using the time domain, frequency domain, and time-frequency domain methods. The features obtained are 68 units with the top 5 features that have the selected RES Index value. The spoken syllables are 21 Thai words with 5 tones for each word. The next research conducted by [7] to extract the EMG signals for Bangli spelling pronunciation using 4 types of methods, namely time-based, frequency based, entropy-based and time-frequency based. The features obtained are 57 units. The syllables are as many as 11 Bangli letters. In this study, we performed facial EMG extraction using discrete wavelet transform to pronounce five syllables in the Acehese language.

III. Method

A. Flowchart research

The stages of this study as shown by the flow diagram in Fig. 1 where the approximate coefficient component and wavelet detail coefficient are obtained by decomposing the wavelet signal on each existing facial EMG data.

B. Data collection

At this stage, the need for the tools needed to retrieve FEMG signals is determined. The tools and equipment are the surface electrode and FEMG sensor as shown in Fig. 2, Arduino Uno mini computers, cables, and laptop.

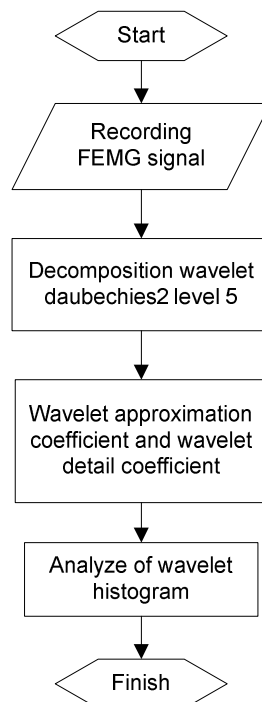


Fig. 1. Flowchart of Research

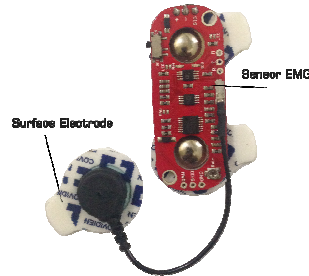


Fig. 2. EMG sensor and surface electrode

The FEMG signal data is recorded and acquired by attaching an EMG surface electrodes sensor to the facial muscles namely the masseter muscle, risorius, and depressor. The electrode placement point for FEMG data retrieval as shown by Fig. 3. Electrodes in the masseter muscle as the reference electrode.

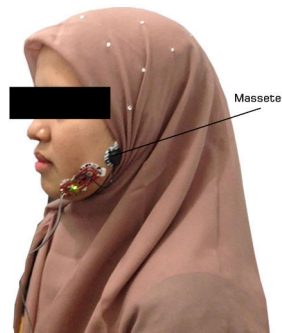


Fig. 3. EMG sensor and surface electrode

The recording of EMG signal uses the Cool Term software. Each word pronunciation is given time for initializing and ending. The initialization and termination times are given for 5 seconds. Initialization time is to normalize the condition of the muscles in the face before speaking the word. While the termination time is to normalize the condition of the muscles in the face after a speech the word. The number of participants involved in taking this data is one person. The five Acehese vocabulary words were spoken as shown in Table 1.

Table 1. Indonesia and Aceh Language Vocabulary

| No | English Language Vocabulary | Aceh Language Vocabulary |
|----|-----------------------------|--------------------------|
| 1. | Where | <i>Ho neuk jak</i> |
| 2. | Go | <i>Ja'</i> |
| 3. | Go home | <i>Ja' woe</i> |
| 4. | Come from | <i>Panè</i> |
| 5. | Who | <i>Soe</i> |

C. Data Processing

After the signal acquisition, the FEMG signal was processed and analyzed using Matlab software. The FEMG signal data was analyzed using the wavelet transform method. The FEMG signal was decomposed first by using two filters namely low pass filter (LPF) and high pass filter (HPF). After decomposition, the FEMG signal was produced two components, namely the approximation coefficient and the detail coefficient. The illustration of the decomposition process is shown in Fig. 4.

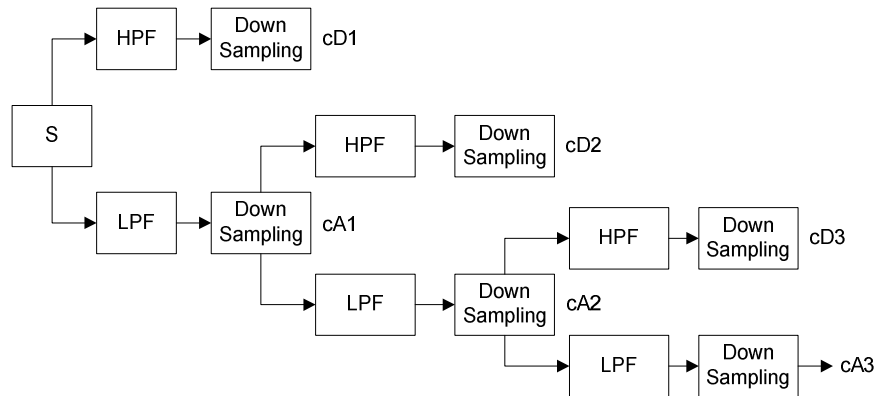


Fig. 4. Illustration of wavelet transform decomposition in EMG signal

Fig. 4 explains that a high pass filter produces a detail coefficient while a low pass filter produces an approximation coefficient. Down sampling is only done at the output of the low pass filter. This down sampling process was repeated using a high pass filter and low pass filter. This research uses the wavelet transform with type Daubechies2 to level 5.

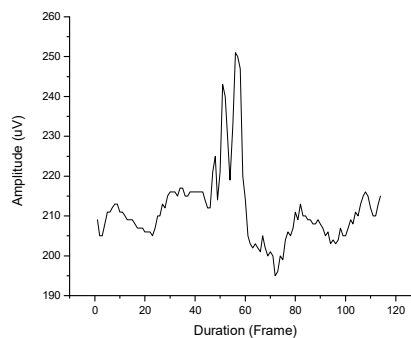
IV. Results and Discussion

Using the discrete wavelet transform method as described in section III above, a facial EMG signal is obtained for pronouncing five Acehese vocabulary words as shown in Table 2.

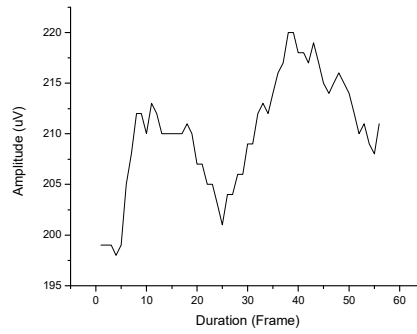
Table 2. Facial EMG signal at the time of the speech

| No | Vocabulary | Facial EMG |
|----|------------|------------|
|----|------------|------------|

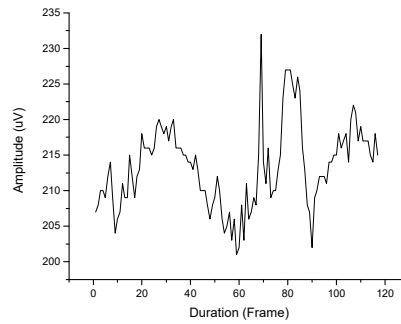
1. *Ho neuk jak*



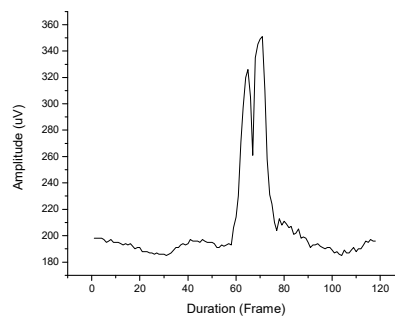
2. *Ja'*



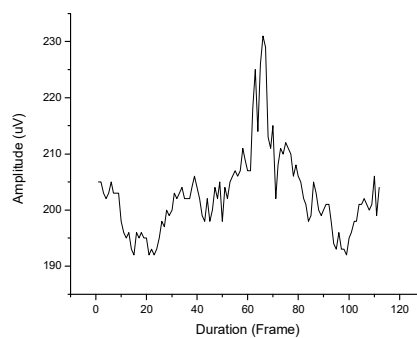
3. *Ja' woe*



4. *Panè*



5. *Soe*

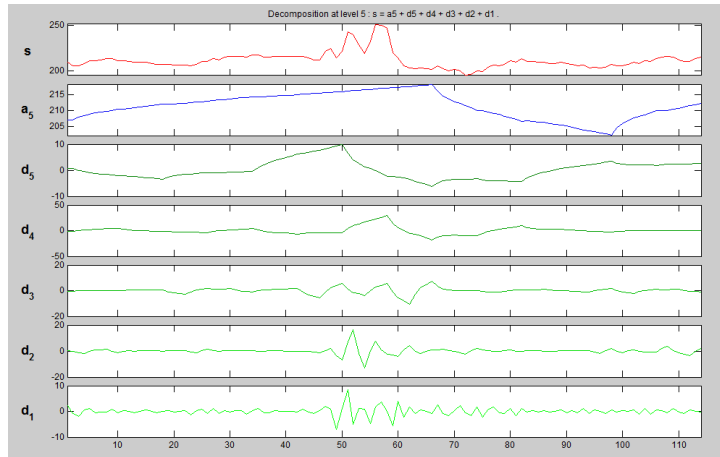


The form of the FEMG signal in the approximation coefficient and the detail coefficient wavelet for the daubechies2 level 5 for each vocabulary word as shown in Table 3.

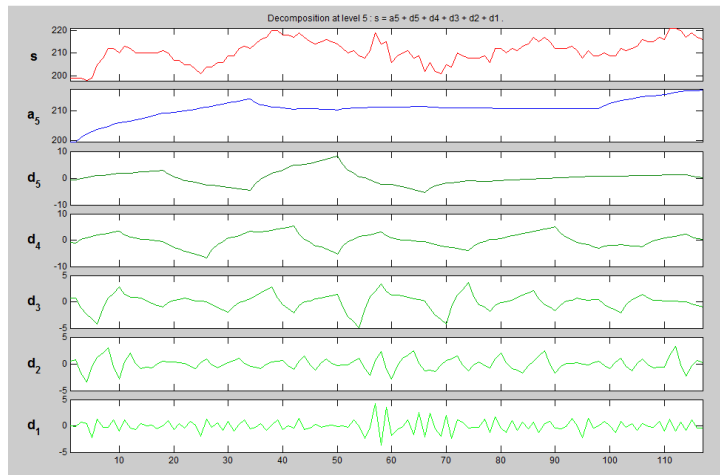
Table 3. The EMG facial signals in the approximate coefficients and detail coefficients wavelet.

| No | Aceh Vocabulary and Wavelet Form |
|----|----------------------------------|
|----|----------------------------------|

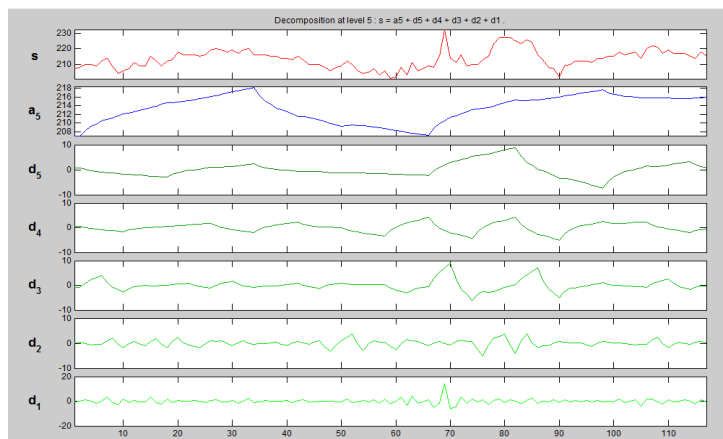
1. *Ho neuk jak* :



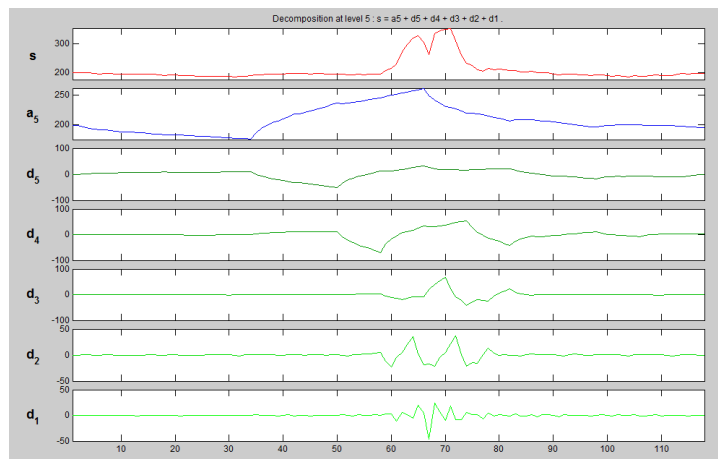
2. *Ja'* :



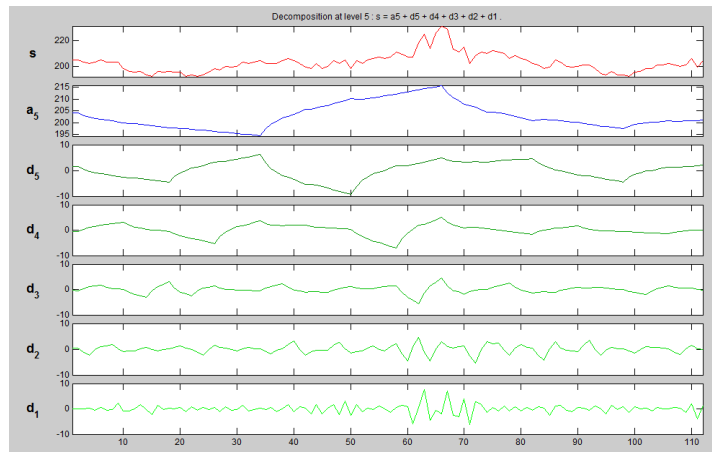
3. *Ja' woe* :



4. *Panè* :



5. *Soe* :



Note : **s** : EMG signal , **a** : approximate coefficient, **d** : detail coefficient

Table 3 shows that the form of the wavelet signal for the approximate coefficients of level 5 is different for each word spoken. Likewise, the shape of the wavelet signal for the Daubechies2 detail coefficient from level 1 to level 5 is also different for each word spoken. This indicates that the signal feature of the FEMG signal has a difference for each spoken word.

The histogram for the approximate coefficient level 1 of the FEMG signal for each word was pronounced as shown in Fig. 5 to Fig. 9.

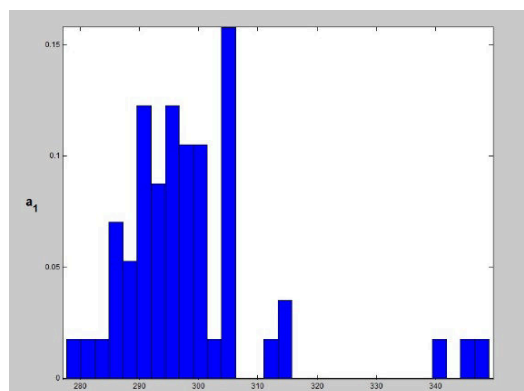


Fig. 5. Histogram of the approximate coefficient for *ho neuk jak*

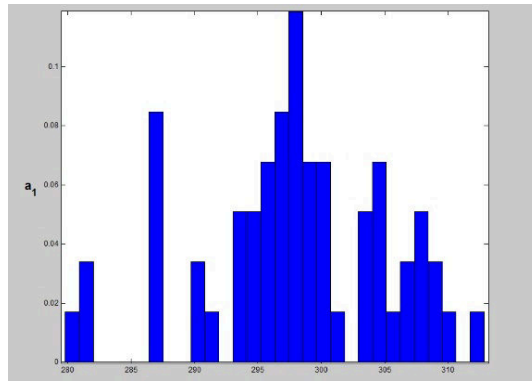


Fig. 6. Histogram of the approximate coefficient for *ja'*

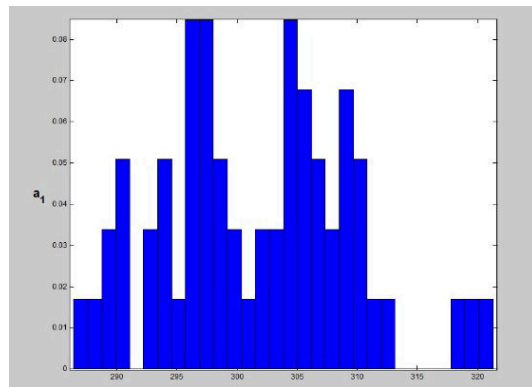


Fig. 7. Histogram of the approximate coefficient for *ja' woe*

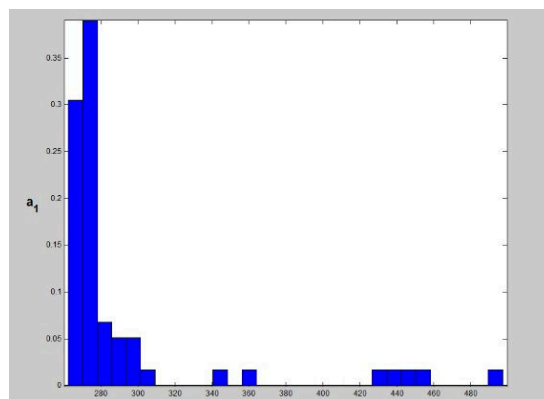


Fig. 8. Histogram of the approximate coefficient for *panè*

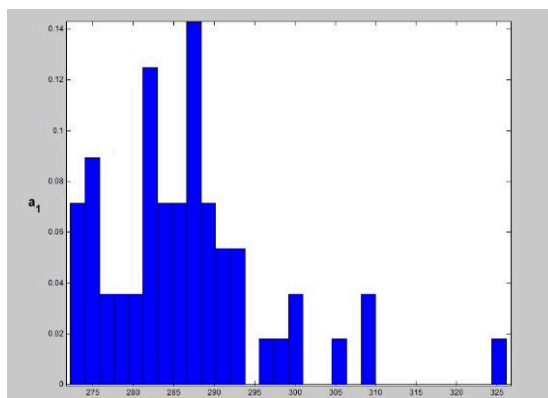


Fig. 9. Histogram of the approximate coefficient for *Soe*

Fig. 5 to Fig. 9 explain that the histogram of approximate coefficient level 1 for each vocabulary word differs the level bar value for each figure according to the word spoken. Fig. 5 shows the maximum level bar with a value of 0.16 in the range of 300 – 310. Fig. 6 shows the maximum level bar with a value of 0.12 at 295 – 300. Fig. 7 shows the maximum level bar with a value of 0.09 in the range of 295 – 300 and 300 – 305. Fig. 8 shows the level bar with a value of 0.37 in the range of 260 – 280. Fig. 9 shows the maximum level bar with a value of 0.14 in the range of 285 – 290. It shows that the FEMG signal for each word pronunciation differently produces a FEMG signal pattern in the risorius, depressor and masseter muscles one another. This EMG signal pattern shows that the humans have different articulation of facial muscle movements for each word spoken so that the resulting FEMG signal features are different from one another. Thus FEMG signal feature can be used to distinguish one another FEMG signal for each word pronunciation.

V. Conclusion

Facial electromyography signals are signal produced because of the contraction of muscles in the human face. The FEMG signal pattern can be known by generating the signal feature. The signal feature is obtained by extracting and analyzing FEMG signals. Extract and analyze of signal feature use the discrete wavelet transform method that aims to determine the level of difference between one another signal at the time of word pronunciation through the approximate coefficient and detail coefficient of the FEMG signal. The results show that the histogram maximum level bar with a value of 0.16 for *ho neuk jak*, 0.12 for *ja'*, 0.09 for *ja' woe*, 0.37 for *panè*, 0.14 for *Soe*.

Acknowledgment

The author's thanks to the Ministry of Research, Technology and Higher Education for the financial support in Beginner Lecturer Research scheme No. D/79.D/POLTAS/IV/2019.

References

- [1] G. L. Read, "Facial Electromyography (EMG)," *Int. Encycl. Commun. Res. Methods*, no. January, pp. 1–10, 2017.
- [2] Y. Golland, A. Hakim, T. Aloni, S. Schaefer, and N. Levit-Binnun, "Affect dynamics of facial EMG during continuous emotional experiences," *Biol. Psychol.*, vol. 139, pp. 47–58, 2018.
- [3] A. Mavratzakis, C. Herbert, and P. Walla, "Emotional facial expressions evoke faster orienting responses, but weaker emotional responses at neural and behavioural levels compared to scenes: A simultaneous EEG and facial EMG study," *Neuroimage*, vol. 124, pp. 931–946, 2016.
- [4] M. Hamed, S. H. Salleh, K. Ismail, A. M. Noor, M. Astaraki, and H. Aslian, "Time-frequency facial gestures EMG analysis using bilinear distribution," *Int. Conf. Signal Image Process. Appl. ICSIPA*, pp. 169–173, 2016.

- [5] Y. Cai, Y. Guo, H. Jiang, and M.-C. Huang, "Machine-learning approaches for recognizing muscle activities involved in facial expressions captured by multi-channels surface electromyogram," *Smart Heal.*, vol. 5–6, pp. 15–25, 2018.
- [6] S. Orguc, H. S. Khurana, K. M. Stankovic, H. S. Leel, and A. P. Chandrakasan, "EMG-based Real Time Facial Gesture Recognition for Stress Monitoring," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 2651–2654.
- [7] S. . Mostafa, M. . Awal, M. Ahmad, and M. . Rashid, "Voiceless Bangla vowel recognition using sEMG signal," *Springerplus*, 2016.
- [8] N. Srisuwan, P. Phukpattaranont, and C. Limsakul, "Three Steps of Neuron Network Classification for EMG-based Thai Tones Speech Recognition," in *International Conference on Electrical Engineering/Electronics, Computer, Telecommunication and Information Technology*, 2013, pp. 0–5.
- [9] M. Lyu, C. Xiong, and Q. Zhang, "Electromyography (EMG)-based Chinese voice command recognition," in *2014 IEEE International Conference on Information and Automation (ICIA)*, 2014, pp. 926–931.
- [10] M. W. Soon, M. I. H. Anuar, M. H. Z. Abidin, A. S. Azaman, and N. M. Noor, "Speech recognition using facial sEMG," in *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, 2017, pp. 1–5.
- [11] Abdussalam, E. H. Rachmawanto, N. A. Setiyanto, D. R. I. M. Setiadi, and C. A. Sari, "Optimasi Keamanan Watermarking pada Daubechies Transform Berbasis Arnold Cat Map," *J. Inform. J. Pengemb. IT*, vol. 04, no. 01, pp. 31–37, 2019.
- [12] E. Lopez-larraz, O. M. Mozos, J. M. Antelis, and J. Minguez, "Syllable-Based Speech Recognition Using EMG," in *International Conference of the IEEE Engineering in Medicine and Biology*, 2010.