Classification of Hand Gestures using sEMG Signals and Hilbert-Huang Transform

Deniz Hande Kisa
Dept. of Biomedical Eng.
Izmir Katip Celebi University
Izmir, Turkey
denizhandekisa@hotmail.com

Mehmet Akif Ozdemir
Dept. of Biomedical Eng.
Izmir Katip Celebi University
Izmir, Turkey
makif.ozdemir@ikcu.edu.tr

Onan Guren
Dept. of Biomedical Eng.
Izmir Katip Celebi University
Izmir, Turkey
onan.guren@ikcu.edu.tr

Aydin Akan
Dept. of Electrical and Electronics Eng.
Izmir University of Economics
Izmir, Turkey
akan.aydin@ieu.edu.tr

***Abstract*—Artificial intelligence is effectively utilized for hand gesture classification in myoelectric systems. In this study, hand movement classification is performed with ML algorithms using electromyography (EMG) signals of 7 hand gestures. The Hilbert-Huang Transform (HHT) was applied to the preprocessed EMG signals to obtain the Hilbert-Huang spectrum (HHS). Six different Gray Level Co-occurrence Matrix (GLCM)-based features were extracted from HHS images. In order to validate the proposed method, the same features were extracted from the snapshots of EMG signals and intrinsic mode functions (IMF) extracted by empirical mode decomposition (EMD), separately. These features are classified with 29 different Machine learning (ML) approaches in the MATLAB® environment. Among these three approaches, the HHS-based novel method yielded the best performance, with an accuracy of 90.87% from the Cubic Support Vector Machine (SVM). The novel HHS and GLCM-based approach may be used in EMG-based biomedical systems as a promising alternative.**

***Index Terms*—Electromyography, GLCM, time-frequency analysis, machine learning, EMD.**

I. Introduction

A surface electromyogram (sEMG) is a non-invasive, inexpensive, and effective tool for electrical muscle activity to control movement assistive devices and rehabilitate physically disabled people [1]. Especially, hand gesture classification from sEMG signals takes an important place in the implementation of myoelectric-based biomedical systems, such as hand prostheses, exoskeletons, sign language, and virtual reality [2]. Artificial intelligence (AI) approaches have shown significant success in the recognition of hand gestures from sEMG signals. Machine learning (ML)-based methods stand out from the other AI approaches regarding their applicability, especially in the myoelectric-based human-machine interface (HMI) applications designed for personal usage [3], [4]. ML methods have the potential to achieve the desired success with less computation time and small datasets that require low hardware resources. One should note that all advantages of ML methods depend on the selected feature for representing the sEMG signals.

The distinct patterns in sEMG signal related to the specific hand gestures can be represented accurately using different parameters. Obtaining such parameters, namely the feature extraction step directly affects the performance of the selected ML method. Variety of features extracted in the time domain (TD) [5], frequency domain (FD) [6], both TD and FD [7], and also in the joint time-frequency (TF) domain [8] have yielded significant performance in ML-based hand gesture classification applications. Besides, the studies in this field continue to explore well-defined muscles for sEMG electrode placement, ideal features, and also ML models with low complexity to enhance the effectiveness of myoelectric devices.

Gray Level Co-occurrence Matrix (GLCM) method, which computes the occurrence of gray-level intensities in neighboring pixels, has been used as a statistical approach to extract texture features from biomedical images [9]. Extraction and analysis of features from various image-based methods by using GLCM and texture-based approaches provide benefits in many areas from disease diagnosis or detection to classification of various biomedical data [9], [10]. In the literature, there are some examples of effective features derived from GLCM like variance, energy, contrast, correlation, etc.

In this study, the GLCM method is used to represent gesture-related sEMG signals in an informative way to improve the performance of ML in myoelectric-controlled systems. First, a TF representation (TFR) is implemented by using the Hilbert-Huang Transform (HHT), which ensures instantaneous specifications in terms of frequency and amplitude in the joint TF domain. HHT provides more localized instantaneous frequency values than traditional TFA methods like Short-Time Fourier Transform (STFT), Wavelet Transform (WT) [11]. Then, GLCM features are extracted from the high-resolution TFR image obtained by the HHS method.

This work was supported by the Izmir Katip Celebi University Scientific Research Projects Coordination Unit [grant numbers 2021-ÖDL-MÜMF-0004, 2022-GAP-MÜMF-0001]; and the Scientific and Technical Research Council of Turkey (TUBITAK) [grant number 120E512].

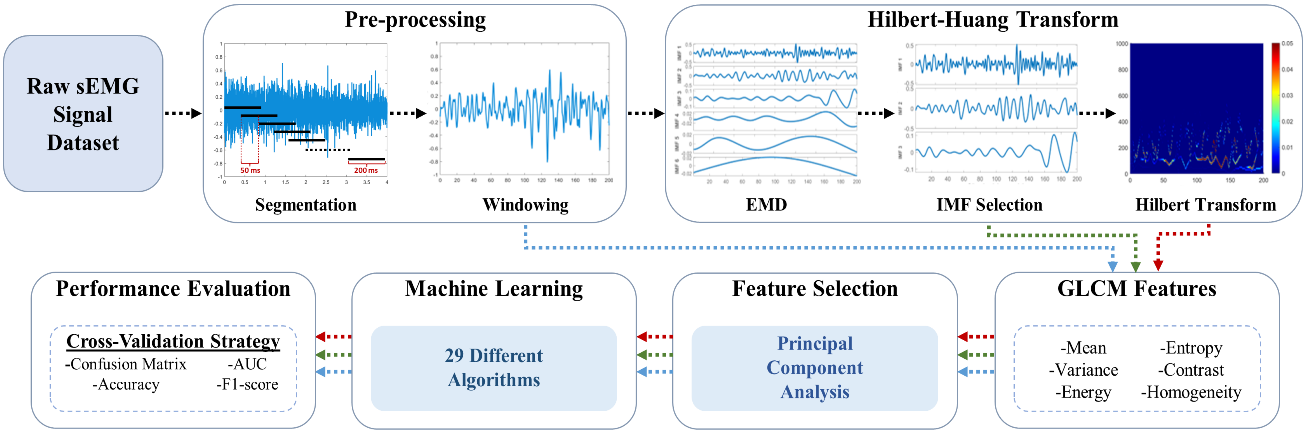

Fig. 1. The framework of this study.

Various well-known ML methods were applied to investigate the performance of the proposed feature extraction approach.

II. Related Work

In this section, we review some remarkable studies that use sEMG and ML-based methods for hand gesture classification to improve the performance of myoelectric-based systems. Arozi et al. [7] performed the pattern recognition study for 9 finger movements. They applied principal component analysis (PCA) to reduce the feature set and yielded an average accuracy of 86.70%. Rabin et al. [8] performed the classification of 6 movements applying the k-nearest neighbors algorithm (kNN) on the STFT features after implementing the PCA, and obtained an accuracy of 77.30%. Devaraj et al. [4] yielded the accuracy of 93.00% using TD features of seven movements with the kNN algorithm. Benalcazar et al. [3] obtained an accuracy of 86.00% with kNN for 5 movements. Sattar et al. [5] performed the recognition of 4 movements via the TD feature, and yielded an average accuracy of 90.70% with the kNN algorithm. Using TD and FD attributes together or separately ensure good classification performance. This study aims to present a potentially successful approach that best represents relevant information by using the features extracted from TFRs that contain both time and frequency information of the underlying signal.

III. Methods

This study proposes a novel GLCM-based feature extraction method from the 3D Hilbert-Huang spectra of sEMG signals for the classification of hand gestures. Well-known ML methods are used to classify the HHS-based features. Also, the proposed method is compared to the features extracted from the snapshots of EMG signals and the IMFs. An illustration of the methodological strategy is shown in Fig. 1.

A. Dataset

In this study, sEMG signals dataset presented in [12] were used. EMG signals are recorded from 4 surface bipolar electrodes, which represent the approximate position of 4 surface muscles (extensor carpi radialis, extensor carpi ulnaris, flexor carpi radialis, and flexor carpi ulnaris). The data was collected from the right hands of 30 healthy subjects and were recorded at a sampling frequency of 2 kHz. Each participant performed 7 specific hand gestures (Fig. 2), which are rest, wrist extension, wrist flexion, wrist radial deviation and ulnar deviation, punch, and open hand in five repetitions. Each gesture was performed in 6 s.

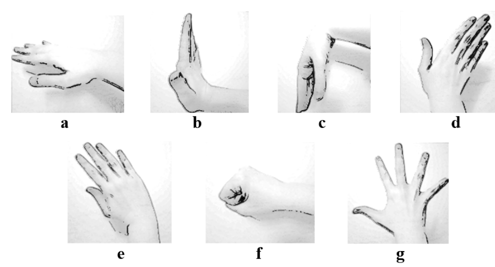

Fig. 2. The used hand movements: a) rest, b) extension, c) flexion, d) radial deviation, e) ulnar deviation, f) punch, and g) open hand.

B. Pre-processing

In order to clear noise from the environment, such as internal organs, electrical devices, or neighboring muscles, a digital bandpass filter with a pass-band of 50-500 Hz and a notch filter at 50 Hz were applied to the recorded sEMG signals. Then, the signals were segmented into 4 s partitions, which represent the steady-states of EMG signals, to eliminate 1 s transient states at the beginning and end [1]. Then, windowing was applied to the 4 s signals by a 200 ms sliding window with the 50 ms increment. As a result, 77 EMG segments were obtained from a 4 s signal. In total, 11550 segments (30 $subjects$ x 5 $repetitions$ x 77 $segments$) were obtained for each gesture.

C. Hilbert-Huang Transform

HHT is an adaptive signal decomposition technique that is used for the processing of non-stationary and non-linear signals like EMG. The analysis of EMG with HHT provides an effective result by specifying motor units and eliminating noise from outside of muscles to extract appropriate EMG features. HHT is composed of two basic steps: empirical mode decomposition (EMD), and the Hilbert Transform (HT).

1) Emprical Mode Decomposition: The EMG signals are adaptively decomposed into a finite number of free intrinsic mode functions (IMF) as $IMF_i(t)$, i = $1, 2, \cdots, N$, where N is the number of intrinsic modes. As such, $x(t)$ may be represented as;

$$x(t) = \sum_{i=1}^{N} IMF_i(t) + r(t) \tag{1}$$

where $r(t)$ is the residue of the signal, from which no IMF can be extracted [13].

2) Selection of IMFs: The most critical step in HHT is the selection of the optimal IMFs to represent the signal. The statistical significance-based ($t-test$) method is conducted in our experiments. In the $t - test$ method, the h-value, which specifies whether the data distribution is normal, and the p-value, which specifies the statistical significance of data, are utilized. The threshold was chosen as $p = 0.05$. The p and h-values of each IMF were calculated, and it was aimed to use IMFs with the greatest p-values. Therefore, IMFs with the three highest p-values were selected.

3) Hilbert Transform: The Hilbert transform was performed using a combination of selected IMFs. The Hilbert transform of the signal $x(t)$ is described as follows:

$$H[x(t)] = y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau \tag{2}$$

where P specifies the Cauchy principal value. An analytical signal $z(t)$ may be obtained from $y(t)$ as:

$$z(t) = x(t) + iy(t) = a(t)e^{j\theta(t)} \tag{3}$$

The instantaneous amplitude ($a(t)$), instantaneous phase ($\theta(t)$) and instantaneous frequency ($\omega(t) = d\theta(t)/dt$) are instantaneous characteristics of the signal. Using the HT, the original EMG segment may be represented as follows (4):

$$x(t) = \sum_{j=1}^{n} a_j(t) e^{i \int \omega_j(t) dt} \tag{4}$$

where n denotes the number of selected IMFs. By the instantaneous features of HT, amplitude and instantaneous frequency may be obtained as functions of time $H(\omega, t)$. HHS provides a 3D representation that contains $a(t)$, $\theta(t)$ and $\omega(t)$ information of the analyzed signal as a TFR [14].

D. Feature Extraction

In ML applications, feature extraction aims to represent the signal in the most informative minimal form. TFR analysis provides distinctive information about the signal. However, the classification of TFR images with an AI method requires an advanced hardware system. To overcome this issue, different approaches have been proposed to represent TFR images in a simple and reduced dimension form. In our study, we use GLCM-based attributes [9] calculated from TFR images and the snapshots of EMGs and IMFs for the classification. Hence, the second-order histogram-based statistical variables of GLCM were computed. These variables are defined as follows:

1) *Mean:* μ gives the average intensity level. Assume that the pixel value $P_{(i,j)}$ at the point (i, j) of an image of size $M \times N$, GLCM mean value is:

$$\mu = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} iP_{(i,j)} \tag{5}$$

2) *Variance:* σ^2 describes the changing intensities around the GLCM mean.

$$\sigma^2 = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_{(i,j)}(i-\mu)^2 \tag{6}$$

3) *Energy:* En denotes the energy contained in the GLCM matrix.

$$En = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_{(i,j)}^2 \tag{7}$$

4) *Entropy:* E is defined as a measure of how much disorder or randomness present in an image.

$$E = -\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_{(i,j)} \log_2(P_{(i,j)}) \tag{8}$$

5) *Contrast:* C specifies the measurement of the drastic alteration in gray level among neighboring pixels.

$$C = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_{(i,j)}(i-j)^2 \tag{9}$$

6) *Homogeneity:* H represents the likeness in gray level among neighboring pixels.

$$H = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \frac{P_{(i,j)}}{1+(i-j)^2} \tag{10}$$

HHT generates 3D energy distributions, however, to extract GLCM-based features, those HHS were projected into 2D images as shown in Fig. 1. The above-explained features are calculated and combined to construct a feature vector. The same features were also calculated by using their snapshots of the EMGs and the selected IMFs extracted by channel-wise EMD [5], [15].

E. Recognition

In order to exhibit the performance of the used feature extraction approach, various ML methods were used to classify seven hand gestures. All classification algorithms were implemented in MATLAB® (R2021a) Classification Learner App (CLA), which includes 29 different ML techniques like SVM, kNN, Random Forest, neural networks (NN), Naïve Bayes (NB), etc. All ML methods in the CLA were tested with all feature sets, but only the methods that achieved the best performance with each set were considered in the study. PCA method was performed on six extracted features before the classification step to represent all features with the lowest variance. Hence, it provided the reduction of the dimension of the predictor space. PCA principle was performed explained in [6]. After obtaining the low-variance data with reduced dimensions, accuracy, F1-score, and area under the receiver operating characteristic curve (AUC) values [1], [16] were calculated to evaluate the performance of the proposed models.

IV. Results and Discussion

In this study, three different feature extraction approaches were assessed according to the classification performance in ML methods. A total of 80850 snapshots were obtained from EMG signal segments and the first three IMFs, and HHS separately. Then, the GLCM-based features were computed by using these images. In the created feature matrix, each EMG channels have 6 GLCM-based features and there are 24 columns in total. Hence, each feature was represented channel-wise. After the implementation of dimensionality reduction, five features are determined by the PCA except for the mean. 5-fold cross-validation technique was used to prevent over-fitting. All computations were executed with an Intel Core i5-6200U CPU at 2.30 GHz. The processing time was calculated as 5.3730 (EMD) + 4.0490 (HHT) + 27.7560 (HHT mesh + Feature Extraction) = 37.1780 x 4 Channel = 148.7120 ms.

TABLE I
Performances of The Best ML Models for Five Features.

Approach	Model	ACC (%)	F1 (%)	AUC
EMG Snapshots	Random Forest	71.34	70.89	0.78
IMF Snapshots	Bagged Trees	82.13	82.74	0.90
HHS Images	Cubic SVM	90.87	90.84	~ 1

The best performance results for three feature extraction approaches were given in Table I. The GLCM-based HHS features achieved the highest ACCy 90.87% with the Cubic SVM method. The ACCy values of punch, flexion, open hand, and ulnar deviation were obtained above 92%, whereas other gestures have ACCy values above 80%. In addition, an error analysis between gestures was performed using the confusion matrix in Fig. 3. The misclassification rate is higher between extension and flexion, radial deviation and ulnar deviation, extension, and radial deviation, and rest and flexion than in other combinations. The reason for this may be related to the similar contraction while performing the different gestures [2]. For instance, when the hand is during the radial deviation, the hand may sometimes be in a form similar to the extension due to the anatomical structure. This may have caused a similar contraction in the muscles. Hence, the distinguishment between HHS of these gestures may reduce due to this issue. Similarly, tend of EMG activity of radial and ulnar deviation was approximated, and muscle activation was similar in the related channels and amplitude levels. This may have caused the misclassification of these gestures. Rather than this misclassification, the overall performance is remarkable. The average F1-score was calculated above 90.84%. When the AUC values were examined, the mean AUC was ~ 1 for the Cubic SVM. Precision, recall, and specificity of the model were obtained as 91.14%, 90.87%, and 98.48%, respectively.

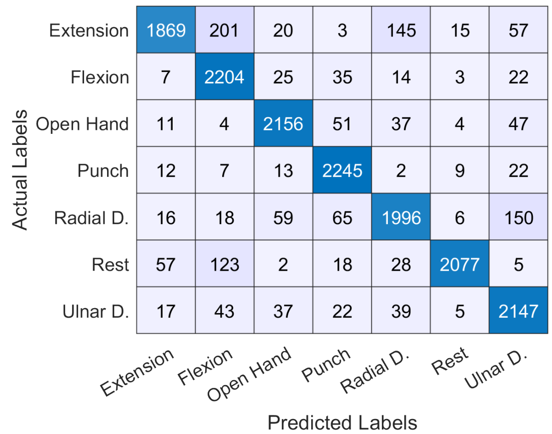

Fig. 3. Confusion matrix of Cubic SVM and HHT-based approach.

The EMG and IMF snapshot-based approaches were behind the HHT-based approach in all tested ML methods. In EMG signal-based approach, the best validation ACCy was obtained as 71.34% by Random Forest. The F1-score was 70.89% and the mean AUC was 0.78. In the IMF-based approach, the best validation ACCy was obtained as 82.13% by Bagged Trees with five features. The F1-score was 82.74% and the mean AUC was 0.90. Here, the feature extraction approach by using the first-three IMFs increased the 10.79% of the classification performance compared to the signal-based approach. As we compare the performance metrics of all the approaches given in Table I, we notice that the TFR-based approach outperforms the time-domain EEG signal and IMF-based approaches [19]. Besides, using the GLCM of TFR for feature extraction has been shown to be an effective approach for hand gesture classification with sEMG signals. The GLCM-based HHS features preserve the information acquired from HHT that determines the intrinsic attributes of short EMG signals, signal frequency characteristics, and also the changing of FD characteristics as time [20].

To assess the effectiveness of HHS-based image features on classification performance, we summarized the performance results of some recent studies published in the literature in Table II. The study in [17] applied binary tree (BT) based SVM method to classify 13 movements. They collected the data from more channels and fewer subjects compared to the presented study. Despite using 10-fold cross-validation, they achieved a 2.67% lower ACC than our best model. Similarly, [8] obtained 13.57% less ACC with STFT-based methods. The multi-channel study in [5], which performed 10-fold cross-validation and classified fewer movements, obtained $\sim 0.17\%$ less ACC than the HHS-based method with half the number of subjects. Also, the study in [15] used HHT and Artificial Neural Network (ANN) for movement recognition but, they used IMF_1-based TD features and yielded the average ACCy of 86.20%. Similarly, the study in [3], which used the dynamic time warping (DTW) algorithm and more channels obtained low ACC compared to this study. Some studies have reported more successful results on the classification. Devaraj et al. [4] and Mahmood et al. [16] used TD features in kNN

TABLE II
COMPARISON TABLE WITH RECENT ML STUDIES.

Study	Year	Mov.	Ch.	Part.	Model	Method	Val.	ACC (%)	F1(%)
[15]	2016	6	2	10	ANN	IMF_1-based features	TTS	86.20	N/A
[17]	2017	13	16	8	BT-SVM	TD features	10-fold	88.20	N/A
[6]	2019	9	16	1	GRNN	TD and FD features	TTS	95.10	N/A
[4]	2020	7	8	32	kNN	TD features	TTS	93.00	N/A
[8]	2020	6	2	5	kNN	STFT-based	10-fold	77.30	N/A
[16]	2021	18	8	3	Fine kNN	TD features	TTS	98.90	N/A
[18]	2021	3	3	5	TKEO+Subspace kNN	TD features	TTS	96.67	N/A
[5]	2021	4	8	15	kNN	TD features	10-fold	90.70	N/A
This study	**2022**	**7**	**4**	**30**	**Cubic SVM**	**HHT features**	**5-fold**	**90.87**	**90.84**

*Mov.: Movement, Ch.: Channel, Part.: Participant, Val.: Validation, TTS: Train-Test Split.

method and achieved a classification ACCy of 93.00% and 98.90% ACCy, respectively. Qi et al. [6] used a combination of TD and FD features and reached 95.10% ACCy with the GRNN method. In another study [18], the combination of Teager–Kaiser energy operator (TKEO) and kNN was used for EMG data classification with TD features. Some of these reported studies suffer from some limitations even though they have reported remarkable results. Some of them used an unbalanced dataset, others datasets suffer from subject-biased, and some of them consider only a few gestures. As a result, we conclude that the proposed method may be an effective alternative approach for feature extraction from sEMG signals in hand gesture recognition.

V. CONCLUSION

This study presents an alternative feature extraction approach based on image features obtained from GLCM of HHS images, which have not been used in hand gesture recognition for EMG-based AI systems before, as far as we know. By comparing the classification accuracy of two different feature extraction approaches, we demonstrate that the GLCM-based approach is able to preserve the pattern embedded in the TFR of the sEMG signal related to the hand gesture. Considering the reducing computational load in the training and prediction of the ML models, we conclude that GLCM-based feature extraction from HHS images has remarkable success in the classification of hand gestures for myoelectric-based devices.

REFERENCES

[1] A. Jaramillo-Yánez, M. E. Benalcázar, and E. Mena-Maldonado, "Real-time hand gesture recognition using surface electromyography and machine learning: A systematic literature review," *Sensors*, vol. 20, no. 9, p. 2467, 2020.

[2] M. A. Ozdemir, D. H. Kisa, O. Guren, A. Akan, and A. Onan, "EMG based Hand Gesture Recognition using Deep Learning," in *Medical Technologies National Congress (TIPTEKNO)*, 2020, pp. 1–4.

[3] M. E. Benalcázar, A. G. Jaramillo, A. Zea, A. Páez, V. H. Andaluz *et al.*, "Hand gesture recognition using machine learning and the myo armband," in *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 1040–1044.

[4] A. Devaraj and A. K. Nair, "Hand gesture signal classification using machine learning," in *2020 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2020, pp. 0390–0394.

[5] N. Y. Sattar, Z. Kausar, S. A. Usama, U. Farooq, and U. S. Khan, "Emg based control of transhumeral prosthesis using machine learning algorithms," *International Journal of Control, Automation and Systems*, pp. 1–11, 2021.

[6] J. Qi, G. Jiang, G. Li, Y. Sun, and B. Tao, "Surface emg hand gesture recognition system based on pca and grnn," *Neural Computing and Applications*, vol. 32, no. 10, pp. 6343–6351, 2020.

[7] M. Arozi, W. Caesarendra, M. Ariyanto, M. Munadi, J. D. Setiawan, and A. Glowacz, "Pattern recognition of single-channel semg signal using pca and ann method to classify nine hand movements," *Symmetry*, vol. 12, no. 4, p. 541, 2020.

[8] N. Rabin, M. Kahlon, S. Malayev, and A. Ratnovsky, "Classification of human hand movements based on emg signals using nonlinear dimensionality reduction and data fusion techniques," *Expert Systems with Applications*, vol. 149, p. 113281, 2020.

[9] D. Kumar *et al.*, "Feature extraction and selection of kidney ultrasound images using glcm and pca," *Procedia Computer Science*, vol. 167, pp. 1722–1731, 2020.

[10] M. A. Ozdemir, G. D. Ozdemir, and O. Guren, "Classification of covid-19 electrocardiograms by using hexaxial feature mapping and deep learning," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1–20, 2021.

[11] M. A. Ozdemir, D. H. Kisa, O. Guren, and A. Akan, "Hand gesture classification using time–frequency images and transfer learning based on cnn," *Biomedical Signal Processing and Control*, vol. 77, p. 103787, 2022.

[12] M. A. Ozdemir, D. H. Kısa, O. Guren, and A. Akan, "Dataset for multi-channel surface electromyography (semg) signals of hand gestures," *Data in Brief*, vol. 41, p. 107921, 2022.

[13] D. H. Kisa, M. A. Ozdemir, O. Guren, and A. Akan, "EMG based Hand Gesture Classification using Empirical Mode Decomposition Time-Series and Deep Learning," in *Medical Technologies National Congress (TIPTEKNO)*, 2020, pp. 1–4.

[14] R. Fonseca-Pinto, "A new tool for nonstationary and nonlinear signals: The hilbert-huang transform in biomedical applications," *Biomedical Engineering, Trends in Electronics, Communications and Software. Croacia: INTECH*, pp. 481–504, 2011.

[15] A. Kukker, R. Sharma, and H. Malik, "Forearm Movements Classification of EMG Signals Using Hilbert Huang Transform and Artificial Neural Networks," in *IEEE 7th Power India International Conference (PIICON)-Conference Proceedings*. Bikaner, India, 2016, pp. 0–5.

[16] N. T. Mahmood, M. H. Al-Muifraje, S. K. Salih, and T. R. Saeed, "Pattern recognition of composite motions based on emg signal via machine learning," *Engineering and Technology Journal*, vol. 39, no. 2A, pp. 295–305, 2021.

[17] H. Huang, T. Li, C. Bruschini, C. Enz, V. M. Koch, J. Justiz, and C. Antfolk, "Emg pattern recognition using decomposition techniques for constructing multiclass classifiers," in *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 2016, pp. 1296–1301.

[18] P. W. Laksono, T. Kitamura, J. Muguro, K. Matsushita, M. Sasaki, and M. S. Amri bin Suhaimi, "Minimum mapping from emg signals at human elbow and shoulder movements into two dof upper-limb robot with machine learning," *Machines*, vol. 9, no. 3, p. 56, 2021.

[19] K. Englehart, B. Hudgins, P. A. Parker, and M. Stevenson, "Classification of the myoelectric signal using time-frequency based representations," *Medical engineering & physics*, vol. 21, no. 6-7, pp. 431–438, 1999.

[20] K. Li, J.-Y. Hogrel, J. Duchêne, and D. J. Hewson, "Analysis of fatigue and tremor during sustained maximal grip contractions using hilbert-huang transformation," *Medical engineering & physics*, vol. 34, no. 7, pp. 832–840, 2012.