A new dataset for the detection of hand movements based on the SEMG signal

1st Adilbek Turgunov
Information technology
Karshi branch of the Tashkent University of Information Technologies named after Muhammad Al-Khwarizmi
Karshi, Uzbekistan
zaxirov91@mail.ru

2nd Kudratjon Zohirov
Computer Systems
Tashkent University of Information Technologies named after Muhammad Al-Khwarizmi
Tashkent, Uzbekistan
qzohirov@gmail.com

3rd Bobur Muhtorov
Psychiatry, Narcology, Child Psychiatry, Medical Psychology, and Psychotherapy
Tashkent Pediatric Medical Institute
Tashkent, Uzbekistan
b.mukhtorov@gmail.com

***Abstract*— in this article, we would like to present a new dataset (DS-dataset) designed to detect hand movements based on SEMG (surface electromyography) signal. This DS includes data from 42 healthy people and seven hand movements, which included three complete arm movements, i.e. punch, grip, finger touch, open hand, three-finger movements, i.e. flexion of the index finger, flexion of the middle finger, flexion of the ring finger, and one waiting state. This data was obtained using BTS's state-of-the-art Free-EMG 10-channel recorder. Based on the data in DS, the characteristic vector of the signal was generated, and were classified using classical classification algorithms (support vector machine - SVM, random forest - RF and k-nearest neighbor algorithm - k-NN). The presented DS can be used as a basis for determining the localization of electrodes and for detecting hand movements when receiving the SEMG correctly.**

***Keywords*— SEMG, dataset, movement recognition, hand, human-machine interface, electrode, classification.**

I. INTRODUCTION

Surface electromyography (SEMG) is a method used to monitor and analyze muscle activity. This signal is widely used in body prostheses, brain-computer interface, and robotics. The collection of EMG signal through low noise has evolved over the last decade.

Many methods and techniques have been developed for the analysis and observation of muscle movement. These methods include electromyography (EMG), mechanomyography, and electroencephalography. Surface electromyography (SEMG) is one of the most advanced methods used today in hand prostheses [1]. SEMG signals have also been widely used in control artificial body parts using muscle activity, in separating noises from signals in human-computer communication (ICA) and brain-computer interface (MKI) systems, and detection of diseases using the heart muscle [2,3], and in studying human movements on an exoskeleton basis [4]. It can be possible for the SEMG signal to record the potentials using electrodes in a non-invasive method (Figure 1).

Surface electrodes can record not only the full activity of the muscle but also the interaction potentials of hundreds or even thousands of nerve fibers and this of course leads to a certain distortion of the needed signal. EMG signals can be recorded by invasive and non-invasive methods. Obtaining EMG signals directly from muscles is called an invasive method. Signal recording by adhering electrodes to the skin surface is called a (SEMG) non-invasive method. The SEMG method has significant advantages over the invasive method, including the fact that obtaining the SEMG signal through electrodes that are easily installed on the skin also provides convenience for the subject. The true appearance of the SEMG signal (raw signal - RS) can be in the range of 0-1000 Hz frequency and amplitude of 0-2 mV. However, the band frequency, which contains important data of the SEMG signal, is in the range of 20-500 Hz. The RS SEMG signal has a very low amplitude, so it can be effected quickly and easily by various noisy environments.

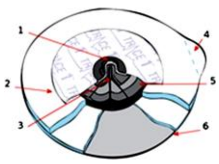

Fig. 1. Electrode structure. Metal corpus (1), plastic cover (2), Ag/AgCl sensor (3), wing (4), electrolyte gel (5), adhesive protective film (6)

Skin resistance reduces the amplitude of the SEMG signal and can cause noise. The noises are generated by effects such as fat between the muscles and the skin, blood flow in the blood vessels. Also, there may be noise from the environment, electromagnetic radiation during recording, magnetic fields propagating from electric power wires, and the external noise. The main noises affecting the SEMG signal occur under the influence of tissue layers between the skin, the zone of innervation (distribution zone - IZ), the intersection of adjacent muscles, the number of electrodes, and their condition. The location of the electrodes can significantly confuse the description of the statistical and spectral factor of SEMG, thus affecting the evaluation of SEMG [5].

There are now several open DSs designed to detect some hand movements. The information about the DSs that have the opportunity for free-to-use is given in Table 1. The most common DS is NinaPro (adapted to non-invasive). The NinaPro DB2 is designed to detect 40 people and 52 gestures.

Most of the many DSs that are open and free do not have the number of repetitions of hand movements, which is very important for the development of recognition algorithms.

2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT) | 978-1-7281-7386-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/AICT50176.2020.9368735

Besides, the sequential repetition of each movement several times (referred to as repetitive in Table 1) may increase the accuracy of the classification of the repetition, but exaggerating the repetitions may also reduce the reliability of the accuracy. This shows its negative effects when we apply the prosthesis to the hand. However, in some DSs, the movements were performed in a random order (shown in Table 1 as sequential and random).

TABLE I. THE MOST COMMON DSS FOR DETECTING HAND MOVEMENTS

Dataset Name	No. of Participants	No. of Gestures	Repetitions	Session	Session Organization	Trials	Gesture
CSL-HDEMG [6]	5	27	10	5	different days	sequential	gesture: 3 s idle: 3 s
EMG Dataset 1 [7]	11	8	12	6	5 poses	sequential	gesture: 5 s idle: 3-5s
NinaPro DB2 [8]	40	49	6	1	-	repetitive	gesture: 5 s idle: 3 s
EMG Dataset 2 [9]	8	15	12	3	-	sequential	gesture: 20 s idle: 3 s
IEE EMG [9]	4	17	32	1	-	sequential	no idle phase
NinaPro DB1 [11]	10	7	12	10	2 per day, 5 days	repetitive	gesture: 8 s idle: 4 s
Megane Pro [12]	10	15	12	10	2 per day, 5 days	repetitive	gesture: 8 s idle: 4 s
NinaPro DB2 [13]	10	52	6	1	-	repetitive	gesture: 5 s idle: 3 s
NinaPro DB3 [13]	10	52	6	1	-	repetitive	gesture: 5 s idle: 3 s
CapgMyo (DB-b) [14]	10	8	10	2	1 day	sequential	gesture: 3 s idle: 7 s
Our dataset (this work)	42	7	30	2	2 week	sequential repetitive	gesture: 3s idle: 7 s

Both approaches increase the reliability of the accuracy rather than performing actions the same, so it is important to perform a large number of repetitions.

The NinaPro DS4 offers a lot of repetition of movements - each movement is repeated 32 times, but only taken from four subjects. Most SEMG datasets only include tests recorded during a single session. This may also not ensure the reliability of the accuracy. Because muscles can be physiologically altered or fatigued, so that it is recommended to receive the signal for as long as possible. In addition, the multiplicity of subjects can lead to changes in the localization of the electrodes, in this situation it is required to pay attention to the physiological structure of the muscles and the innervation zone.

Thus, DSs involving multiple sessions are of great positive importance. Four of the DSs given in Table 1 (NinaPro DB1, Megane Pro, CSL-HDEMG, CapgMyo) were performed over multiple sessions and daily intervals. SEMG signals were obtained in real-time mode in the DS given in Table 1.

In this work, the signal was recorded using six electrodes. The DS included six active hand movements, as well as the idle state. In a comparison of existing DSs, the significant advantages were that each movement arose as a result of a large number of repetitions, and data were collected from 42 subjects. It has two repetitions in data collection. All data were recorded within two weeks. The number of repetitions was increased to 30. The duration of the movements is 3 s.

II. ORGANIZATION OF THE DATASET

A. *Experimental Setup*

A special experiment was developed in the organization of this DS. The settings in this process have anatomical features. 10-channel EMG analyzers of the BTS Company were used to receive the SEMG signal. The data is received at a frequency of 1000 Hz and the size of the digital converter is 12 bits. Besides, a filter was used to eliminate interference. Data were obtained in a non-invasive method using Ag-Cl adhesive electrodes. Six electrodes were used to receive the signal. The electrodes were installed in innervation zones.

B. *Experimental Design and Participants*

Data on 7 movements were collected in the formation of the DS (Figure 2). The action selection was based on previous research and was selected as the main action for the SEMG-based HMI. DS consists of 6 active movements (hand punch, open hand, hand grip, and grasping movements with three fingers) and one idle state.

As a result, each participant repeats each movement 20 times. Each session was held for at least a week, and in the active state, the number of repetitions was increased to 30.

Each subject participated in the experiment on a voluntary and voluntary psychiatric basis. Before the starting the experiment, participants were asked about their gender, age, weight, health history, tobacco activity, and sports activities. In addition, the diameters of the wrists and elbows were measured. The experiment was performed on 42 healthy people including 5 women and 37 men aged 19 to 37 years. Each subject underwent the experiment 2 times.

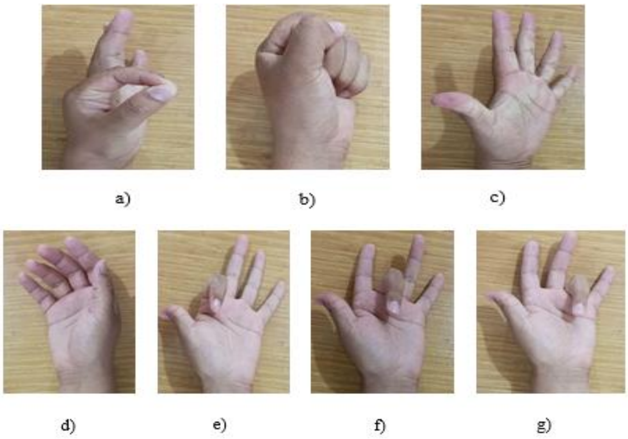

Fig. 2. Used hand movements. (a - touch finger, b - hand punch, c - open hand, d - idle, e - flexion the index finger, f - flexion the middle finger, g - flexion the ring finger)

C. *Pre-Processing*

The obtained SEMG signals were filtered (20-500 Hz). These filtered signals were used to calculate the features used in the section above. The calculations were performed using windows with a duration of 256 ms. The dataset of each subject was divided into separate groups and processed separately.

III. FEATURE EXTRACTION AND CLASSIFICATION

The experiences of using the modern classification methods as a basis for the organization and future application of the required feature vector of the DS that we have established were described in this section.

A. *Signal amplitude*

The SEMG signal is highly variable, the main reasons for this are the resistance of the human skin, the quality of the electrodes, the location of the tendon and innervation zones, and so on. Signal filtering, segmentation, and normalization works are done to reduce the above problems in many scientific types of research. In this work, we expressed the signal amplitudes in idle states of the SEMG amplitudes as the signal-noise ratio (SNR). This allows a direct comparison of signals during inactive and active movements, which is very important for classification. The formula of the SNR is expressed as follows:

$$SNR = 20 * \log\left(\frac{P_{gesture}}{P_{idle}}\right) \quad (1)$$

where, $P_{gesture}$ the average signal strength of the active movement, P_{idle} refers to the average signal strength in the idle state.

The SNR results of the average of each movement in an average of five motion channels in each movement are shown in Figure 3. The movement of the hand punch showed higher activity ($p < 0.05$) than the other movements (average 15 dB). Punch and handgrip movements were higher than 11 dB, the SNR was much lower in the touch finger movement of the hand, and the SNR dropped to an average of 3-6 dB in the downward movement of the hand.

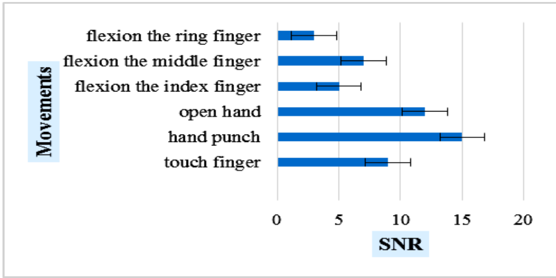

Fig. 3. SNR analysis of movements. All values are calculated for idle time, which means that in this case it is considered noise.

B. *Feature Extraction*

The advantages of this DS differ from others due to many factors, for example, the DS has many subjects and many repetitive movements. DS can be used to classify movements to distinguish the constant features of an object.

So far, many scientific types of research have been conducted on the separation and classification of EMG signal features. In this work, we used SSI, ACC, IEMG, WL, RMS, and MAV features that we used in our previous scientific researches [15]. These features provided an accuracy of up to 99%, but in our previous works, a high result was achieved due to the low number of movements (3) and the number of subjects (20). Even if this set of features we used is calculated in a time zone, they represent the signal amplitude, frequency, and complexity. The classification of movements was performed using three different classifiers (RF, k-NN, SVM).

C. *Classification*

In this section, the comparisons between different combinations of classifiers and feature sets are provided. The accuracy was calculated for each class and based on their average weight.

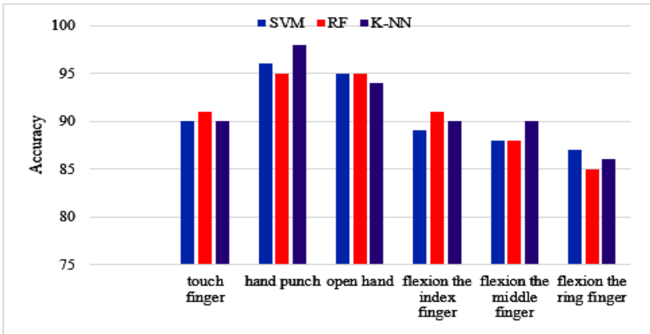

Fig. 4. Classification results

Tests have shown that the RF, k-NN, and SVM algorithms in the feature sets have all achieved good results. The difference between them was mainly noted between hand and finger movements (Figure 4).

As can be seen from figure 4, the best result was hand punching and gave 96%, 95%, 98% results, respectively, in

SVM, RF, k-NN algorithms. The relatively lowest result was flexion of the ring finger movement (87%, 85%, 86%, respectively).

IV. DISCUSSION

Noise will be 7-8 dB less than other hand movements compared to finger movements. This means that the higher the noise, the greater the involvement of the muscles in the movement.

However, in NinaPro DS, no significant amplitude differences of movements were detected [13]. The main reason for this is that the elbow (standing on the table, tilting the elbow) does not move when the hand is moving. There is no significant difference in amplitudes due to the lack of noise, the accuracy is very high. However, great accuracy does not justify itself when testing the hardware later. In the DS we are proposing, there is no obstruction to the elbow during the data recording process, the elbow is in a free position.

Let us explain another difference; this is an experimental procedure and data processing. Unlike other DSs, our DS has a much higher noise level because the elbow moves freely. By determining the noise level, we will be able to determine the filtering limits. We limit the data in all DSs according to its noise level, not the same filtering. There is also a significant difference between the finger movements and other movements and in their amplitudes or the level of noise in the signal. This condition is observed because of the muscles less contraction than the full movement of the hand when the fingers move individually.

The second important difference is the number of subjects and the experimental order. More than 15 people did not participate in many scientific types of research. It should be noted that the excessive number of the class of movements also leads to a decrease in inaccuracy. For example, it is around 40-50 in the NinaPro DS, but the accuracy rate is 74% [13].

V. CONCLUSIONS

In this article, 7 movements and one idle state of the hand are classified, and a total of 42 people participated in this experiment. The important aspects were focused on in the DS we purposed such as the location of electrodes, preparation of the dataset for classification, first determining the noise level for filtering and setting the filter limit accordingly, correct selection of signal characteristics to increase classification efficiency.

Using the most modern classification methods, the classification of data was achieved with an accuracy of 85% to 98%. Although this is a negative result in some scientific works, the organization of the dataset is characterized by the fact that it consists of data from different participants. It should be noted that the main purpose of this article was to present a new DS and to show new approaches in the organization of the dataset. The results obtained have shown that our DS can be used in the future as a basis in the detection of hand movements and in the application of this work in hand prostheses or in the HCI systems.

REFERENCES

[1] Weir JP. Mechanomyographic and electromyographic responses during fatigue in humans: influence of muscle length. European J Applied Physiol 2000; 81: 352-359.

[2] B. Muminov, R. Nasimov, S. Mirzahalilov, N. Sayfullaeva and N. Gadoyboyeva, "Localization and Classification of Myocardial Infarction Based on Artificial Neural Network," 2020 Information Communication Technologies Conference (ICTC), Nanjing, China, 2020, pp. 245-249, doi: 10.1109/ICTC49638.2020.9123300

[3] Muminov Bakhodir Boltaevich, Nasimov Rashid Hamid ogli, Gadoyboyeva Nigora Soibjon qizi and Mirzahalilov Sanjar Serkabay ogli. "Estimation affects of formats and resizing process to the accuracy of convolutional neural network." 2019 International Conference on Information Science and Communications Technologies (ICISCT) (2019): 1-5. doi: 10.1109/ICISCT47635.2019.9011858

[4] Kiguchi K, Tanaka T, Fukuda T. Neuro-fuzzy control of a robotic exoskeleton with EMG signals. IEEE Transact Fuzzy Syst 2004; 12: 481-490.

[5] Cescon C, Raimondi EE, Zaest V, Drusany-Stari K, Martsidis K. Characterization of the motor units of the external anal sphincter in pregnant women with multichannel surface EMG. Int Urogynecol J 2014; 25: 1097-1103.

[6] Amma, C.; Krings, T.; Böer, J.; Schultz, T. Advancing muscle-computer interfaces with high-density electromyography. In Proceedings of the Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; ACM: New York, NY, USA, 2015; pp. 929–938.

[7] Khushaba, R.N.; Takruri, M.; Miro, J.V.; Kodagoda, S. Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features. Neural Netw. 2014, 55, 42–58.

[8] Atzori, M.; Gijsberts, A.; Castellini, C.; Caputo, B.; Hager, A.G.M.; Elsig, S.; Giatsidis, G.; Bassetto, F.; Müller, H. Electromyography data for non-invasive naturally-controlled robotic hand prostheses. Sci. Data 2014, 1, 140053.

[9] Khushaba, R.N.; Kodagoda, S. Electromyogram (EMG) feature reduction using mutual components analysis for multifunction prosthetic fingers control. In Proceedings of the 2012 12th International Conference on Control Automation Robotics & Vision (ICARCV), Guangzhou, China, 5–7 December 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 1534–1539.

[10] Cene, V.H.; Tosin, M.; Machado, J.; Balbinot, A. Open database for accurate upper-limb intent detection using electromyography and reliable extreme learning machines. Sensors 2019, 19, 1864.

[11] Palermo, F.; Cognolato, M.; Gijsberts, A.; Müller, H.; Caputo, B.; Atzori, M. Repeatability of grasp recognition for robotic hand prosthesis control based on sEMG data. In Proceedings of the 2017 International Conference on Rehabilitation Robotics (ICORR), London, UK, 17–20 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1154–1159.

[12] Giordaniello, F.; Cognolato, M.; Graziani, M.; Gijsberts, A.; Gregori, V.; Saetta, G.; Hager, A.G.M.; Tiengo, C.; Bassetto, F.; Brugger, P.; others. Megane pro: Myo-electricity, visual and gaze tracking data acquisitions to improve hand prosthetics. In Proceedings of the 2017 International Conference on Rehabilitation Robotics (ICORR), London, UK, 17–20 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1148–1153.

[13] Pizzolato, S.; Tagliapietra, L.; Cognolato, M.; Reggiani, M.; Müller, H.; Atzori, M. Comparison of six electromyography acquisition setups on hand movement classification tasks. PLoS ONE 2017, 12, e0186132.

[14] Phinyomark, A.; Scheme, E. EMG pattern recognition in the era of big data and deep learning. Big Data Cogn. Comput. 2018, 2, 21.

[15] A. Turgunov, K. Zohirov, A. Ganiyev and B. Sharopova, "Defining the Features of EMG Signals on the Forearm of the Hand Using SVM, RF, k-NN Classification Algorithms," 2020 Information Communication Technologies Conference (ICTC), Nanjing, China, 2020, pp. 260-264, doi:10.1109/ICTC49638.2020.9123