الجمهورية الجزائرية الديمقراطية الشعبية
**People's Democratic Republic of Algeria**
وزارة التعليم العالي والبحث العلمي
**Ministry of Higher Education and scientific research**
المدرسة الوطنية العليا للتكنولوجيا
**National Higher School of Technology**
Department of Electrical Engineering and Industrial Computing

**Final Year Project to Obtain the Diploma of Master**

Field
**Automation**

Speciality
**Automation and Industrial Computing**

- Subject -

# Gesture Recognition Using Surface Electromyography and Deep Learning for Prostheses Hand: State-of-the-Art

By
**ZITOUNI Nour Elislam** **SERAICHE Oubada**

**Examination Committee :**

| | | |
|---|---|---|
| **M. Cherabit Nourdine** | Chair | M.C.B |
| **Ms. REBAI Karima** | Supervisor | M.C.A |
| **M. Choutri Kheireddine** | Examiner | M.C.B |
| **Ms. Chouaf Seloua** | Examiner | M.A.A |

**Algiers, the 04/04/2023**

**Academic year 2022 –2023**

# Gesture Recognition Using Surface Electromyography and Deep Learning for Prostheses Hand: State-of-the-Art

line 1: 1st Given Name Surname
line 2: *dept. name of organization (of Affiliation)*
line 3: *name of organization (of Affiliation)*
line 4: City, Country
line 5: email address or ORCID

line 1: 2nd Given Name Surname
line 2: *dept. name of organization (of Affiliation)*
line 3: *name of organization (of Affiliation)*
line 4: City, Country
line 5: email address or ORCID

line 1: 3rd Given Name Surname
line 2: *dept. name of organization (of Affiliation)*
line 3: *name of organization (of Affiliation)*
line 4: City, Country
line 5: email address or ORCID

***Abstract*— Limb amputation reduces the patient's ability to perform daily living tasks. It is possible to handle this loss of autonomy by restoring certain functionalities of the amputated limb using a myoelectric prosthesis. This device performs various movements according to the muscular electrical activity of its user. It has been shown that a large majority of amputees can voluntarily generate muscle activation patterns that differ according to the movement they wish to perform. In addition, for actuation and reproducible movements of the upper limb, these activation patterns are also reproducible. They are thus a reliable and intuitive way to generate movements with a prosthesis. This is the best approach to providing the natural control wanted by these amputees.**

**The development of a new method for the control of a myoelectric prosthesis involves the combination of several disciplines, such as human physiology, signal processing, and machine learning.**

**Deep learning (DL) based solution has been used successfully in the development of intelligent systems in recent years and can effectively improve the quality of myoelectric prosthesis. It can be used for gestures recognition from surface Electromyogram (sEMG)**

**The objective of this study is to compare different deep learning architectures for gesture recognition using surface electromyography (sEMG) signals in myoelectric prostheses. The aim is to identify the most effective architecture or combination of architectures to improve control and functionality for individuals with limb amputation.**

***Keywords*— *Surface EMG, Convolution neural network Transfer learning, Sensor fusion, hand gesture recognition, features extraction, prosthesis hand, recurrent neural network***

## I. INTRODUCTION

Biomedical signals are essential for describing physiological phenomena and are critical for evaluating and diagnosing various pathologies. For example, signals like EMG, EEG, and ECG can be used for this purpose. Specifically, in this work we focus on the application of Electromyography (EMG), Which is a technique used to measure and record the electrical activity of muscles. EMG plays a significant role in studying, assessing, and diagnosing neuromuscular disorders. Additionally, this signal is also useful for controlling prosthetic hands or arms.

To understand how amputees are able to produce myoelectric signals useful for the control of a prosthesis, it is necessary to know how the human body copes with the loss of a member. After an amputation, the amputee's body tries to adapt to the severed tissue and damaged nerves. It is not uncommon that after this neuromuscular rearrangement, the amputee has the impression that their limb is still present after the operation. This phenomenon is known under the name phantom limb and can last a lifetime, although it tends to fade with time[1].

A 1998 study found that amputees may experience a variety of sensations from of their phantom limb [2]. These sensations include muscle spasms, changes in temperatures or even pain. In addition, a majority of amputees are able to perform from the voluntary movements of their phantom limb (PLM : phantom limb movements) to repetition. Amputees report that the sensation of movement and the way the limb ghost moves are different from before their amputation. PLM also produces a surface EMG (sEMG) signal which is different from the signal normally observed in a healthy subject, hence the importance conducts trials with amputees. Each amputee has activation patterns unique that depend on how their wound was closed and reorganized neuromuscular. The sEMG pattern will be substantially similar for each repetition of the same physical task or movement, but will be different for other movements. Additionally, these sEMG patterns are reproducible from one session to another during data acquisition. Therefore, the recognition of sEMG patterns and the corresponding gestures can be used for the control of artificial prosthesis.

A gesture recognition system that uses EMG and deep learning, consists of three stages: data acquisition, pre-processing, and classification. Multiple electrodes can be used to acquire EMG data from muscles, but using too many channels can lead to unnecessary complexity and increased computational power, without significantly affecting the system's design [3].

The classification consists of determining to which class the feature extracted from a EMG belongs. Classical Machine Learning algorithms do not have the ability to analyze a complete sEMG signal for make their prediction. For sEMG classification, it is usually recommended to extract and use relevant features instead of using all the available data. The sEMG signals is first divided into time windows before extracting features on each window in each acquired channel

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

because instantaneous values of the sEMG signal are usually useless for the classification algorithm due to their stochastic and sequential nature [4][5].

In the last years, different approaches are developed to tackle this problem of classification and recognition of limbs motions (gestures) using classical technics [2][5][6]. However, the evolution of Deep Learning (DL) algorithms and their performances that surpassed conventional ML methods in several areas, a large number of researchers proposed new DL-based approaches.

The purpose of this paper is to review this category of approaches and to study some examples of recently developed methods. To this end, we start by presenting the characteristics of the sEMG signals. Then, we discuss recent works that addressed the problem of sEMG classification for gestures recognition.

Regardless of the architecture employed, the main purpose is to improve the classification accuracy and the classifier's robustness, and one of the main factors in obtaining an accurate prediction is the amount of training data available.

Although the data obtained from healthy individuals and amputees is not exactly the same, there is a significant decrease in the classification accuracy for amputees. The majority of algorithms have been trained using data from healthy participants, which might allow for easier and larger recruitment of participants in studies using healthy participants.

Additionally, there is a difference between offline and online classification results, in other words, the performance of the classifier in real-time cannot be predicted through the use of a static data set. Many publications propose offline models [3][6][7] and obtain high accuracy (90–96%), but only a few propose a real-time model[7]. Finally, the problem of real-time hand gesture recognition is still open for new methods and approaches.

In this paper, we review the efficiency of DL in gesture recognition based on sEMG signals and some technique include data acquisition, data pre-processing, feature extraction, pattern classification, post-processing. Also, some classification methods and approaches.

## II. ELECTROMYOGRAPHY

Electromyography (EMG) is a technique used to measure the electrical activity generated by muscle contractions during neuromuscular activities caused by voluntary movements. The electrical signals produced by these muscle contractions (Fig. 1) can be acquired through EMG and can be used to control prosthetic devices. EMG is a valuable tool in the field of neuromuscular research and has contributed significantly to advancements in prosthetics and rehabilitation.

Carrying out the everyday activities, such as walking, standing, or lifting objects, requires the simultaneous contraction of several muscles. In turn, these muscle contractions are the result of central nervous system (CNS) control. The CNS typically controls posture and motion through motor unit (MU) activation. These MUs represent the minimal functional structures in muscles that generate electrical activity. This electrical activity can be detected by intramuscular (iEMG) or surface electromyography (sEMG) techniques [9].

EMG signals reflect the intention and control of the patient's movement and are produced by the contraction of muscle fibers, which generate electrical impulses that can be detected and analyzed with special equipment. In this work, we are interested on the application of (sEMG) for the recognition of hand gestures. These signals are commonly used in the design of prosthetic hands to provide a means of controlling the movement of prosthetic devices [9]. In this application, EMG sensors are placed on the residual limb, and signals from muscle contractions are used to control the movement of the prosthetic hand. The EMG signals are processed and translated into commands that control the movement of the prosthetic hand, enabling the user to perform various grasping and manipulation tasks. For example, the user can generate an EMG signal by contracting their biceps muscle, which can be used to open or close the prosthetic hand.

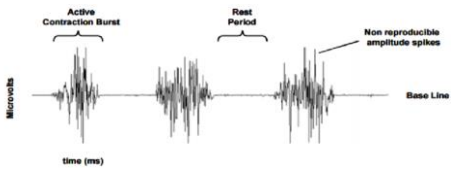



**Fig1.** The electrical signal emitted by the muscles

### A. *Data acquisition*

The measurement of the EMG signal is carried out by means of sensors either invasively (iEMG) or non-invasively (sEMG):

#### 1) *Intramuscular EMG (iEMG)*

This type of EMG measurement uses needle or wire electrodes that are inserted directly into the muscle tissue. They directly affect muscle fibers and make it possible to obtain the signal emitted by a single fiber. However, they have the disadvantage of potentially damaging the skin and muscle tissue.

#### 2) *Surface EMG (sEMG)*

Surface electrodes are fixed onto the skin above the targeted muscle. They provide the signal emitted by a group of muscle fibers, rather than just one as with needle or wire electrodes. However, sEMG signals can be affected by various factors, such as skin impedance, improper electrode placement, and cross-talk from adjacent muscles.

Among surface electrodes, we distinguish:

- Gelled electrodes use a conductive gel as a chemical interface between the skin and metal part. The gel layer is typically made of Ag-AgCl. These electrodes have a limited lifespan and are designed for single use only.



Fig 2. Needle electrode

- Dry electrodes are typically used in situations where the size of the electrode makes it difficult to use a gel. Unlike gelled electrodes, dry electrodes are reusable and have a longer lifespan. However, they have a higher electrical impedance, which means the quality of the sEMG signal may be slightly distorted by sweating.



Fig 3. Surface electrodes

The choice between sEMG and iEMG depends on the specific application and the information required from the EMG signal. sEMG is commonly used in non-invasive applications, such as prosthesis control, while iEMG is typically used in research and clinical applications, such as the diagnosis of neuromuscular disorders.

### *B. EMG electrode placement*

In order to measure accurately, the electrodes must be placed in the correct location on the skin, which is typically over the belly of the muscle being studied. The orientation and spacing of the electrodes can also affect the quality of the recorded signal. Proper skin preparation, such as cleaning and shaving if necessary, is important to reduce noise and ensure good electrode-skin contact. Additionally, the placement and orientation of the electrodes may need to be adjusted for different individuals, as variations in body size, shape, and muscle location can affect the recorded signal.

The preferred location for the electrodes is between the motor point (or innervation zone) and the tendinous insertion, As illustrated in Fig. 4.

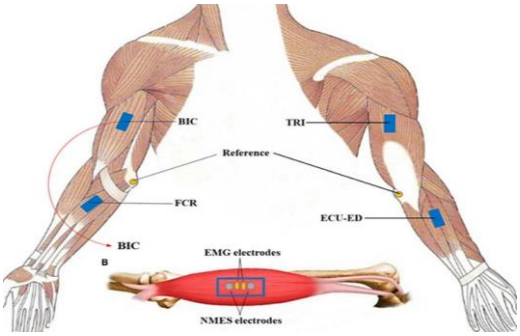



Fig 4. Placement of electrodes

### *C. Pre-processing*

EMG signals are typically noisy, and significant preprocessing is often necessary to extract a useful signal before they can be classified for controlling interfaces as shown in figure 5.

#### *1) Amplification*

The acquired EMG signals are weak and need to be amplified using an amplifier (The amplitude of the EMG detected is 10 mv maximum). The amplifier should have a high input impedance to avoid loading the signal source and a low output impedance to prevent signal degradation due to impedance mismatching. The choice of amplifier and gain adjustment method depends on the specific application and the level of noise in the EMG signal.

#### *2) Signal filtering*

The detection of EMG is easily affected by noise (temperature change, interference between signals emitted by neighboring muscles, etc.) which can affect their accuracy and reliability. So, the signal must therefore be filtered to remove unwanted noise. The filtering process is typically done using band-pass filters, which allow only a certain frequency range of the EMG signal to pass through while attenuating frequencies outside this range.

#### *3) A/D conversion*

sEMG are analog signals. In order to be stored and processed by a computer, they must be sampled and each sample converted to digital form, a binary number. In this way, each sample can be assigned a number.

Overall, the preprocessing steps ensure that the sEMG signals are clean, stable, and normalized, enabling feature extraction and classification algorithms to accurately recognize different gestures.

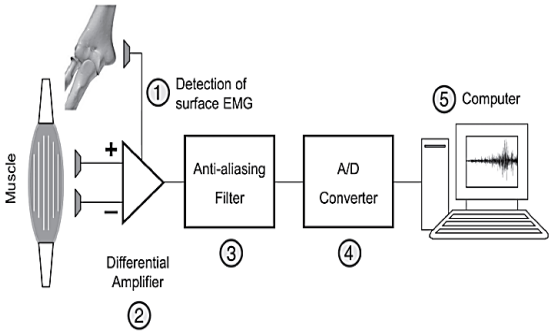



Fig 5. Simplified synoptic diagram of the pre-processing steps of sEMG signal

## III. CLASSIFICATION USING DEEP LEARNING

There are different approaches proposed in the last years for sEMG gestures recognition. In this section, some methodz are presented

- **Convolutional Neural Network (CNN)**

CNN is the first and most deep learning architecture applied to sEMG signal recognition (> 50%) [10]. It is still the most common and popular method due to its well-recognized results and the large number of applications provided in the literature. The results show that CNN can give more precise classification details compared to conventional machine learning methods.

- **CNN combined with handcrafted features**

CNN has been used for image classification; thus, images can be generated from sEMG signals and fed to the CNN block to extract specific features automatically.

The CNN network proposed in [1] contains two convolutional layers, a max pooling layer, and a global average layer to avoid overfitting when using a small dataset (Fig. 6).

The combination of features extracted using CNN and features extracted from time spectral analysis is fed to a multi-layer perceptron classifier (MLPC). This approach reaches an accuracy of 98% with four hand gesture classes of 100 signals each and a 750 ms duration.

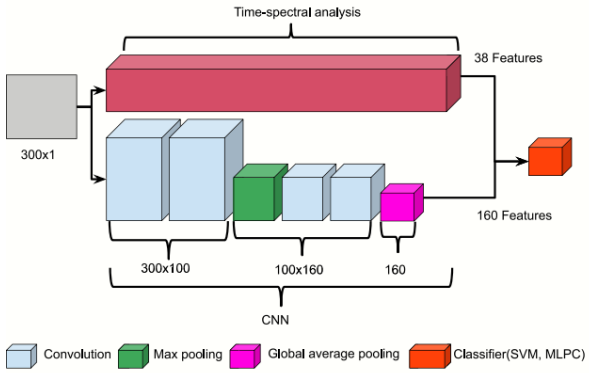



Fig. 6: The proposed combined features architecture [15]

- **CNN with LSTM block**

In [15], the CNN network is used to extract spatial features from sEMG data, and the LSTM (Long and Short Time Memory) block is used to extract temporal features existing in the sEMG time series. This block contains three LSTM layers, the first two of which consist of 32 cells each and the last one of which has 64 cells.

The spatial features extracted from CNN are concatenated with the temporal features obtained from the LSTM block, and the result is fed to a fully connected layer consisting of 100 neurons in the hidden layer and 17 in the final layer representing the 17 gesture classes. This architecture is designed in parallel to increase the computational speed and avoid the latency in the series architecture. This hybrid model achieves an accuracy of 98% according to [12]. Finally, the hybrid techniques have the potential to improve the classification accuracy in this field.

- **CNN combined with RNN**

The combination of CNN and RNN (Reccurent Neural Network) is the most popular structure for sEMG-based applications that have been used by researchers for years [16][17].

A combination of these two networks leads to the extraction of spatial and temporal information at the same time, this combination usually shows better performance than a single network.

Additionally, the pure CNN model achieves an accuracy of 78.7% in 14 able-bodied subjects with 16 independent movements, while the CNN-RNN model reaches an accuracy of 90–99% using different datasets, according to [5].

Finally, this type of architecture can extract spatiotemporal features efficiently.

- **3D-CNN using sEMG and accelerometers data and transfer learning**

In [14], a CNN block is designed based on the LetNet-5 structure, which is adapted to sEMG images. Each convolutional layer contains a ReLu activation function to remove negative values and increase the non-linearity of the network. The first layer consists of 32 filters, and the second has 64 layers, followed by a max-pooling layer to reduce the computation complexity, and then this layer is connected to a fully connected layer with a SoftMax max in the output.

Sensor fusion is a technique of combining data from multiple sensors to improve accuracy when considering some limitations of each individual sensor. Thus, the 3D-CNN [1] is designed to combine information from sEMG signals and accelerometer signals. These signals are recorded simultaneously and form a 3D data frame.

- **Transfer learning**

Transfer learning is a solution to the problem of lack of data needed for the training of deep learning algorithms requires a huge training dataset, which is not available in all cases for a number of reasons, one of which is the lack of participants, especially amputees. Additionally, this seems like a solution to improve accuracy and reduce the need for more data.

In [13], the network is pre-trained with the MNIST dataset and its weights values are saved after training. The weights of the first two layers of this network are fixed with the pre-trained values and the other layers are trained using a fully connected layer with an sEMG dataset.

The accuracy achieved using transfer learning without sensor fusion and only pure sEMG data is about 78%. The use of data fusion has improve this result and an accuracy of 94% is achieved.

- **2D-CNN with Majority voting technique compared to 3D-CNN network**

3D-CNN with 3D kernels is capable of extracting spatial and temporal information from sEMG data, contrary to 2D-CNN, which is limited to extracting spatial information only. Obviously, 3D-CNN would achieve a higher accuracy than the pure 2D-CNN model.

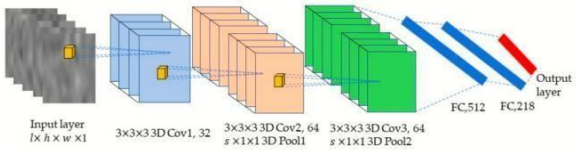



Fig. 7: 3D Convolutional Network

The proposed 2D-CNN network consists of parallel 2D-CNN models that have been trained with the same dataset. The majority voting technique was set to vote over multiple consecutive results to achieve a higher accuracy than the pure 2D-CNN model.

The 3D-CNN and the 2D-CNN+MV results with different sliding windows are illustrated in Fig. 8 [16]:

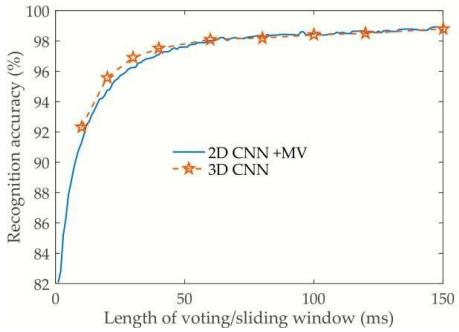



Fig. 8. The comparison of recognition accuracies between 3D CNN over different sliding windows and 2D CNNs over different voting windows.

- **Temporal convolutional network (TCN)**

TCN is a technique that uses the whole raw of the sEMG signal instead of windowing the signal under 300 ms. This technique is used to obtain a higher accuracy of classification with a limited computational resource. Compared to the classic CNN that can automatically learn spatial features, TCN uses a one-dimensional convolutional layer running along the time dimension in order to discover and extract temporal information from a given sEMG input.

TCN achieves higher accuracy with EMG data as input, three healthy participants, and 10 classes[12]. The accuracy reached a value of 93.7% and also overperformed the CNN on 53 classes task [14].

In this work, data is collected using 12 sensors and provided to a one-dimensional convolutional layer with 10 kernels, these features are then concatenated and fed to the first block of the TCN model [8].

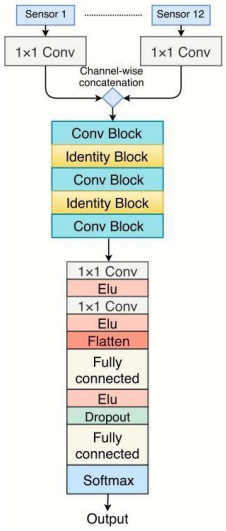



Fig. 9: TCN model [8]

The proposed architecture consists of two main blocks: an identical block to extract temporal features, and a convolutional block consisting of convolutional layers with different kernels. After this process, the data is fed to a fully connected layer with a SoftMax function in the last layer. This TCN model achieves an accuracy of 92.5 %.

## IV. Conclusion

EMG signals are electrical signals that are captured, recorded, and then used in many medical applications. In addition, they can provide more detailed information about the functioning status of human muscles. The exploitation and use of these signals remain important research, in order to produce systems and applications of a medical or functional nature using advanced methods and techniques of signal processing, such as Deep Learning,

The role of deep learning methods in gesture recognition from sEMG signals is to automatically learn to capture distinguishing features of the unique muscle activation patterns associated with different gestures and use these features to make accurate predictions of gesture classes and to establish a robust control for hand prostheses.

The performance of the classification system can be further improved by adjusting the parameters of the classification algorithm or by incorporating additional features or preprocessing techniques.

Overall, the classification of EMG signals is an important step in the analysis of muscle function and can be applied in a variety of applications, including prosthetic control, rehabilitation, medical diagnosis, and sports performance analysis.

## References

[1] M.Z. ur Rehman, A. Waris, S. Gilani, M. Jochumsen, I. Niazi, M. Jamil, D. Farina, and E. Kamavuako. "Multiday EMG-Based Classification of Hand Motions with Deep Learning Techniques." Sensors, vol. 18, no. 8, pp. 2497, Aug. 2018.

[2] E. Rahimian, S. Zabihi, S.F. Atashzar, A. Asif, and A. Mohammadi. "Surface EMG-Based Hand Gesture Recognition via Hybrid and Dilated Deep Neural Network Architectures for Neurorobotic Prostheses." Journal of Medical Robotics Research, vol. 05, no. 01n02, pp. 2041001, Mar. 2020.

[3] Peng, F., Chen, C., Lv, D., Zhang, N., Wang, X., Zhang, X., and Wang, Z. ”Gesture Recognition by Ensemble Extreme Learning Machine Based on Surface Electromyography Signals.” Medical Rehabilitation Research Center, Shandong Institute of Advanced Technology, Chinese Academy of Sciences, Jinan, China.

[4] ] Dezhen Xiong, Daohui Zhang, Xingang Zhao and Yiwen Zhao, ”Deep Learning for EMG-based Human-Machine Interaction: A Review,” IEEE/CAA J. Autom. Sinica, vol. 8, no. 3, pp. 512-533, Mar. 2021.

[5] ] Raez, M.B.I., Hussain, M.S., and Mohd-Yasin, F. ”Techniques of EMG signal analysis: detection, processing, classification and applications”.

[6] U. Cote-Allard, C.L. Fall, A. Drouin, A. Campeau-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,

[7] Boca A, Park DC. Myoelectric signal recognition using fuzzy clustering and artificial neural networks in real time.Boca A, Park DC. Myoelectric signal recognition using fuzzy clustering and artificial neural networks in real time.

[8] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[9] Dezhen Xiong, Daohui Zhang, Xingang Zhao and Yiwen Zhao, ”Deep Learning for EMG-based Human-Machine Interaction: A Review,” IEEE/CAA J. Autom. Sinica, vol. 8, no. 3, pp. 512-533, Mar. 202

[10] Dezhen Xiong, Daohui Zhang, Xingang Zhao and Yiwen Zhao, ”Deep Learning for EMG-based Human-Machine Interaction: A Review,” IEEE/CAA J. Autom. Sinica, vol. 8, no. 3, pp. 512-533, Mar. 2021. doi:

10.1109/JAS.2021.1003865 shu Deep Learning for EMG-based Human-Machine Interaction: A Review

[11] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[12] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[13] Chen, J., Bi, S., Zhang, G., Cao, G. (2020). HighDensity Surface EMG-Based Gesture Recognition Using a 3D Convolutional Neural Network. Sensors

[14] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE

[15] Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[16] X. Gao, M. Iwase, J. Inoue, and E. Maeda, “Poster: Gesture recognition based on ConvLSTM-attention implementation of small data sEMG signals,” in Proc. ACM Int. Joint Conf. Pervasive and Ubiquitous Computing and Proc. ACM Int. Symp. Wearable Computers, London, UK, 2019, pp. 21..24.

[17] B. A. Xie, H. B. Li, and A. Harland, “Movement and gesture recognition using deep learning and wearable-sensor technology,” Proc. Int. Conf. Artificial Intelligence and Pattern Recognition, Beijing, China, 2018,