Near-Infrared Spectroscopy Driven Human-Machine Interface based on Convolutional Neural Networks

A. González, J. Illescas, D. González, M. Pérez, J. Fajardo, and A. Lara
BiomedLab
Universidad Galileo
Guatemala City, Guatemala
{edwin.silva, pablo.rodriguezillescas, duglasa, mperez, julio.fajardo, andrealh}@galileo.edu

***Abstract*—Human-machine interfaces based on the classification of hand gestures often use electromyography to collect and interpret user intent; however, this methodology records the electrical activity produced by skeletal muscles suffering from different drawbacks that affect their performance. In this way, other approaches, such as near-infrared spectroscopy, were studied, presenting advantages over traditional techniques since it is not susceptible to electrical noise and electrode degradation. This work presents a hand gesture classification system for a human-machine interface based on near-infrared spectroscopy and a convolutional neural network that classifies five gestures. The data acquisition system shows a transparent and fluent transfer of data that represents the hemodynamics of three specific sets of muscles used with a classification model for five different hand gestures, showing comparable and promising results against traditional electromyography-based classification methods.**

***Index Terms*—Near-infrared spectroscopy, pattern recognition, hand gesture classification, human-machine interface, convolutional neural networks.**

I. Introduction

The increasing demand for human-machine interfaces (HMI) within fields like robotics and assistive technologies, such as prostheses, has catalyzed extensive research efforts to achieve efficient and robust classification of hand gestures [1]. This effort has enabled the study of diverse methods to recover and interpret user intent [2]–[5]. In this way, hand gesture classification based on surface electromyography (EMG) relies on the electrical activity generated by the muscles during the activation and engagement of hand movements. This technique is widely used because it is minimally invasive and can produce a fair amount of information from the signal about the amount and positioning of the muscle fibers engaged [6], [7]. Thus, several HMI work exclusively by analyzing specific EMG activation profiles or by pattern recognition methodologies based on machine learning techniques [2], [8]. However, the use of EMG for this purpose has well-known drawbacks, such as electrode positioning, fatigue, inherent crosstalk in the surface signal, displacement of the muscles, and the limb position effect [2], [9]–[11].

On the other hand, some efforts have been made to study different types of technologies to monitor the activity of the muscles [12]. These emerging technologies could solve some of the problems present in the techniques based on EMG, especially in interpreting the user's intention from the classification of manual gestures. Human-Machine interfaces based on near-infrared spectroscopy (NIRS) have shown to be a good alternative since it had been demonstrated the relationship between microvascular oxygenation levels and muscle activation [13]–[17]. This way, when specific muscular contractions occur, the perfusion varies according to the hand gesture performed, as well as the geometry and characteristics of the optodes used in the acquisition systems [18]. Thus, the work proposed in [13] performs a selection process that resulted in 11 key features, enhancing gesture recognition accuracy. In addition, comparative testing using different classifiers: multilayer perceptron neural network (MLPNN), Bayesian classifier (BC), linear and quadratic discriminant analysis (LDA and QDA)—was conducted on these features across various gestures and participants. The study found that the MLPNN model presents the most promising result, confirming the chosen features' potential to extract valuable information from hemodynamic signals.

This work presents the proof of concept for an HMI based on a data acquisition system for muscle activity based on the hemodynamic response of three different sets of muscles using NIRS optodes strategically located in a 3D printed bracelet, as shown in Fig. 1. In addition, a Convolutional Neural Network (CNN) model is presented for hand gesture classification for

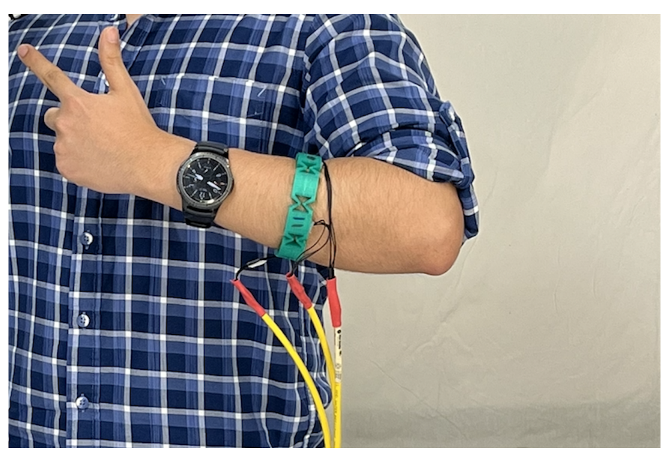

Fig. 1. The NIRS acquisition system placed on the forearm of one of the volunteers. The bracelet facilitates the location of the three optodes in three specific groups of muscles.

five gestures captured using the NIRS data acquisition system, presenting comparable results with a model based on the same architecture but trained with a similar EMG dataset.

The rest of this work is structured as follows: Section II elaborates on the system architecture of the acquisition system and the dataset preparation. Section III, delves into the experiment design, the data preprocessing, and the hand gesture classification model proposed. Additionally, experimental results about the system's functionality are presented in Section IV. Finally, the conclusions are presented in Section V.

II. METHODS

A. Hardware Architecture

The data acquisition was implemented using a high-performance unit (MCU) based on the ARM Cortex-M4F architecture (STM32F303K8), which employs its I2C module for communicating with three optodes. This way, three Maxim Integrated, MAX30102, configured for SPO2 operation, with a sampling rate of about 400 Hz, and an ADC resolution of 18 bit, were used to capture the hemodynamics associated with muscle contractions gathered from three different groups of muscles. Since each optode has the same slave address (fixed and unmodifiable from the factory), a simple time multiplexing method is implemented by providing the bus clock signal of the bus to each optode using an AND gate, as shown in Fig. 2. Thus, data available in the internal FIFO of each MAX30102 is retrieved and filtered within a period of 50 ms to be transmitted and storage into a personal computer (PC) using UART communication protocol with a baud rate of about 230400 bauds.

B. Dataset

NIRS signals were obtained from 10 subjects (8 male, 6 right handed and 2 left handed and 2 right handed females) using 3 optodes attached to the forearm at the flexor digitorum superficialis, extensor digitorum, and flexor carpi ulnaris. Each subject performed 5 hand gestures: Hand Close (HC), Hand Open (HO), Wrist Flexion (WF), Wrist Extension (WE), and Index Finger Extension (IFE), as shown in Fig. 3. Each of these movements was repeated 25 times, having a resting period of 10 seconds between repetitions to avoid muscle fatigue. Finally, 3750 NIRS signals were acquired to train and test the model.

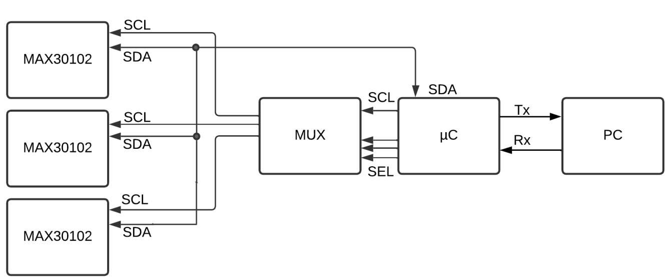

Fig. 2. NIRS system acquisition diagram block for the three optodes used to prepare the dataset.

EMG signals used in the second model were obtained from the Gesture Recognition and Biometrics ElectroMyogram database (GRABMyo) [19]. This dataset was acquired from 43 subjects using surface EMG. To perform the comparison between methods, only the signals from the 5 hand gestures were extracted using the channels where the electrodes were positioned as in NIRS. The final dataset is composed of 4515 signals.

C. Classification model

The classification of the hand gestures using the EMG and NIRS signals was performed using a Convolutional Neural Network (CNN). The model architecture is based on a predefined architecture [20] which consists of 5 blocks as presented in Fig. 4. The first blocks are composed of Convolutional layers and a Rectified Linear Unit (ReLU) as the activation function, and the final layer is a Fully Connected layer with a Softmax activation function. The models were trained using Categorical Cross-entropy as the loss function. Additionally, two dropout layers were added, one after the second block and one after the forth block. The dropout rate was set to 0.8 in both layers.

III. EXPERIMENT DESIGN

A. Signal Pre-processing

Before logging the data in a PC using a comma-separated value (CSV) format, a simple DC blocker filter in cascade with a third-order low pass Butterworth Filter with a cut-off corner frequency of about 10 Hz is implemented to complement the on-chip sensor's discrete-time proprietary filtering (electrical noise and ambient light cancelation). This way, the following IIR filters (coefficients truncated with 4 decimal digits) was implemented in the MCU to remove drift and the noise introduced by the pulsations of the IR light from each optode.

$$H_1(z) = \frac{1 - z^{-1}}{1 - 0.9500z^{-1}} \tag{1}$$

$$H_2(z) = \frac{0.0004 + 0.0012z^{-1} + 0.0012z^{-2} + 0.0004z^{-3}}{1 - 2.6862z^{-1} + 2.4197z^{-2} - 0.7302z^{-3}} \tag{2}$$

Thus, all the filtered data from the three optodes are stored in a small ping-pong buffer of 8 blocks and then sent via UART to the PC data logger.

Fig. 3. Gestures performed by the subjects. From left to right: Hand Open (HO), Wrist Flexion (WF), Wrist Extension (WE), Hand Close (HC), and Index Finger Extension (IFE), respectively.

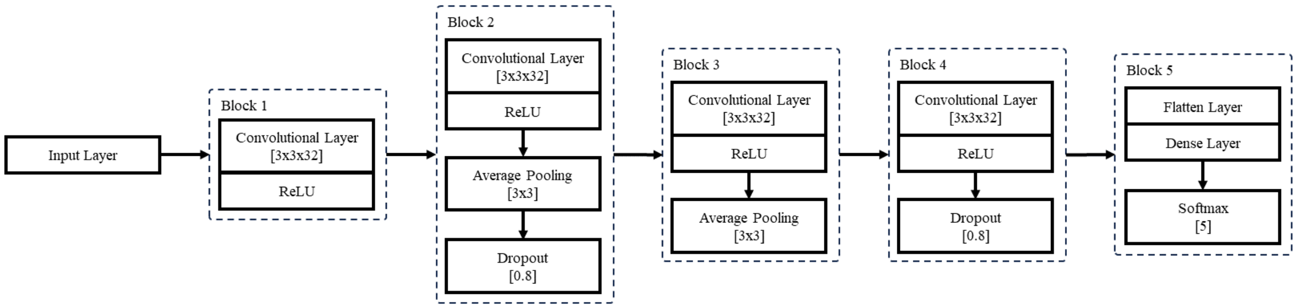

Fig. 4. CNN model architecture used for hand gesture classification for both EMG and NIRS-based methods.

B. *Classification Network*

A CNN architecture model (as shown in Fig. 4) was implemented to classify the five hand gestures captured by the NIRS data acquisition system. Additionally, a second model using the same architecture was trained with EMG data to compare and validate the feasibility of the system. An 8-fold cross-validation approach was used for both models, as well as the ADAM optimizer with a learning rate of 0.001 and a batch size of 512. The EMG-based model was trained for 60 epochs, while the NIRS-based model was trained for 600 due to the reduced amount of data.

IV. Results

The data acquisition system was successfully tested and validated, as shown in the first row of Fig. 5. The results show real-time operation as the time delay introduced due to buffer management, signal pre-processing (Eqs. 1 and 2), and data transmission results in approximately 50 ms. The first row of Fig. 5 presents hemodynamic signals captured from three different hand gestures (WE, IFE, and HO) acquired from the three optodes placed on the sets of muscles mentioned in Sec. II-B at a sample rate of about 400 Hz, and compared with EMG signals acquired with the OTBioelettronica's EMG-USB2+ with a sample rate of about 2048 Hz (GRABMyo dataset) [19]. In addition, the orange, green, and blue signals on Fig. 5 represent muscle activity on the extensor digitorum, flexor digitorum superficias, and flexor carpi ulnaris set of muscles, respectively.

The EMG and NIRS classification models achieved an accuracy of 80 % and 82 %, respectively. As shown in Fig. 6, NIRS signals allowed an overall improved detection of hand gestures compared to EMG. Specifically, hand gestures such as WE, IFE, and HO can be better detected by the NIRS model, suggesting that NIRS signals allow improved gesture detection.

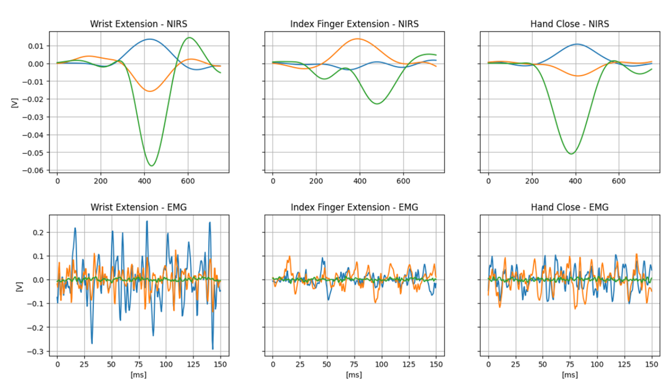

Fig. 5. Comparison between EMG and NIRS raw data using three different gestures, (a) IFE, (b) HO and (c) WF, respectively.

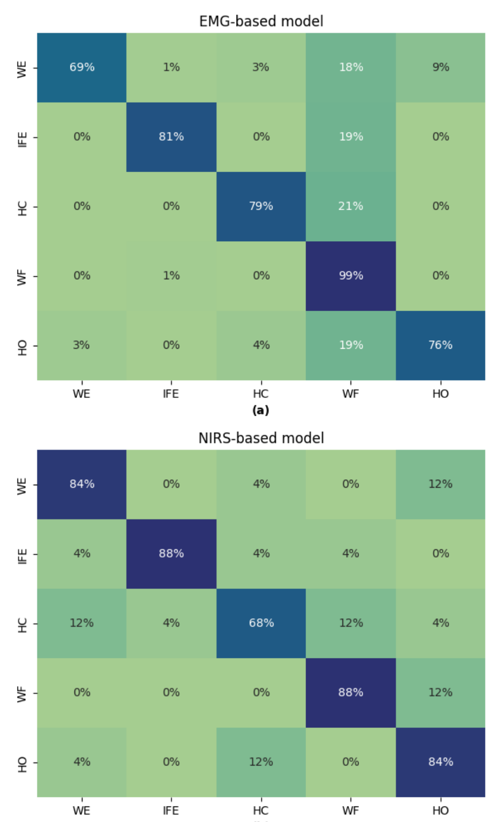

Fig. 6. Confusion Matrices of both models for the testing set. (a) EMG-based model, (b) NIRS-based model.

V. Conclusions

In this work, we presented a data acquisition system based on three NIRS optodes and a hand gesture classification model based on CNN. The results suggest that gesture classification with an accuracy of 82 % can be achieved based on NIRS signals and a CNN, as shown in Figs. 5 and 6. Moreover, the comparison between the EMG and NIRS classification models showed that NIRS signals provide information that allows more accurate detection of gestures where there is similar muscle activation such as WE, IFE, and HO. The latter demonstrates an advantage over EMG signals as the model fails to distinguish between them.

Furthermore, NIRS-based methodologies offer distinct advantages over their EMG counterpart, including hand gesture classification. One well-known advantage is their capability to capture multi-joint movements simultaneously, avoiding the crosstalk effect between an adjacent set of muscles. This benefit enables more comprehensive hand gesture recognition through enhanced spatial coverage, such as those involved in the hand gesture processes (involving the movement of multiple finger joints). On the other hand, muscle fatigue estimation also benefits from this methodology, as it provides non-invasive monitoring of hemodynamic changes in muscle tissue, offering more profound insight into oxygenation and metabolic activity, which are crucial indicators of this process [17].

In addition, NIRS methodologies are also immune to electrical noise, which makes it ideal for use in rehabilitation processes based on functional electrical stimulation (FES) [16].

Finally, in this work, we performed experiments as proof of concept to demonstrate the potential of NIRS-based systems for hand gesture detection. Moreover, we compared it against other traditional methods such as EMG. In the future work, we want to extend the dataset by including more subjects and optodes, and further investigate other CNN architectures to improve hand gesture detection that can help in the development of human-machine interfaces.

References

[1] J. Fajardo, G. Maldonado, D. Cardona, V. Ferman, and E. Rohmer, "Evaluation of user-prosthesis-interfaces for sEMG-based multifunctional prosthetic hands," *Sensors*, vol. 21, no. 21, p. 7088, 2021.

[2] A. Fougner, Ø. Stavdahl, P. J. Kyberd, Y. G. Losier, P. Parker *et al.*, "Control of upper limb prostheses: terminology and proportional myoelectric control a review," *Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 5, pp. 663–677, 2012.

[3] E. Moutopoulou, G. A. Bertos, A. Mablekos-Alexiou, and E. G. Papadopoulos, "Feasibility of a biomechatronic EPP Upper Limb Prosthesis Controller," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 2454–2457.

[4] G. Hotson, D. P. McMullen, M. S. Fifer, M. S. Johannes, K. D. Katyal, M. P. Para, R. Armiger, W. S. Anderson, N. V. Thakor, B. A. Wester *et al.*, "Individual finger control of a modular prosthetic limb using high-density electrocorticography in a human subject," *Journal of neural engineering*, vol. 13, no. 2, p. 026017, 2016.

[5] D. Johansen, C. Cipriani, D. B. Popović, and L. N. Struijk, "Control of a robotic hand using a tongue control system—a prosthesis application," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1368–1376, 2016.

[6] R. H. Chowdhury, M. B. Reaz, M. A. B. M. Ali, A. A. Bakar, K. Chellappan, and T. G. Chang, "Surface electromyography signal processing and classification techniques," *Sensors*, vol. 13, no. 9, pp. 12 431–12 466, 2013.

[7] B. Ahkami, K. Ahmed, A. Thesleff, L. Hargrove, and M. Ortiz-Catalan, "Electromyography-based control of lower limb prostheses: a systematic review," *IEEE Transactions on Medical Robotics and Bionics*, 2023.

[8] G. N. Saridis and T. P. Gootee, "EMG pattern analysis and classification for a prosthetic arm," *IEEE Transactions on Biomedical Engineering*, vol. BME-29, no. 6, pp. 403–412, 1982.

[9] A. Fougner, E. Scheme, A. D. Chan, K. Englehart, and Ø. Stavdahl, "Resolving the limb position effect in myoelectric pattern recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 6, pp. 644–651, 2011.

[10] F. Cordella, A. L. Ciancio, R. Sacchetti, A. Davalli, A. G. Cutti, E. Guglielmelli, and L. Zollo, "Literature review on needs of upper limb prosthesis users," *Frontiers in neuroscience*, vol. 10, p. 209, 2016.

[11] K. S. Türker, "Electromyography: Some Methodological Problems and Issues," *Physical Therapy*, vol. 73, no. 10, pp. 698–710, 10 1993. [Online]. Available: https://doi.org/10.1093/ptj/73.10.698

[12] L. Guo, Z. Lu, and L. Yao, "Human-machine interaction sensing technology based on hand gesture recognition: A review," *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 4, pp. 300–309, 2021.

[13] E. Nsugbe, "A pilot exploration on the use of NIR monitored haemodynamics in gesture recognition for transradial prosthesis control," *Intelligent Systems with Applications*, vol. 9, p. 200045, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S266730532100034X

[14] S. Jiang, P. Kang, X. Song, B. P. Lo, and P. B. Shull, "Emerging wearable interfaces and algorithms for hand gesture recognition: A survey," *IEEE Reviews in Biomedical Engineering*, vol. 15, pp. 85–102, 2022.

[15] J. McIntosh, A. Marzo, and M. Fraser, "SensIR: Detecting hand gestures with a wearable bracelet using infrared transmission and reflection," in *Proceedings of the 30th annual ACM symposium on user interface software and technology*, 2017, pp. 593–597.

[16] M. Yang, Z. Yang, T. Yuan, W. Feng, and P. Wang, "A systemic review of functional near-infrared spectroscopy for stroke: current application and future directions," *Frontiers in neurology*, vol. 10, p. 58, 2019.

[17] B. R. Scott, K. M. Slattery, D. V. Sculley, R. G. Lockie, and B. J. Dascombe, "Reliability of telemetric electromyography and near-infrared spectroscopy during high-intensity resistance exercise," *Journal of Electromyography and Kinesiology*, vol. 24, no. 5, pp. 722–730, 2014.

[18] R. Boushel and C. Piantadosi, "Near-infrared spectroscopy for monitoring muscle oxygenation," *Acta Physiologica Scandinavica*, vol. 168, no. 4, pp. 615–622, 2000. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-201x.2000.00713.x

[19] N. Jiang, A. Pradhan, and J. He, "Gesture recognition and biometrics ElectroMyogram (GRABMyo)." [Online]. Available: https://physionet.org/content/grabmyo/1.0.2/

[20] M. Cognolato, M. Atzori, D. Faccio, C. Tiengo, F. Bassetto, R. Gassert, and H. Muller, "Hand gesture classification in transradial amputees using the myo armband classifier," in *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*. IEEE, 2018, pp. 156–161.