2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Actor-Critic Reinforcement Learning Based Algorithm for Contaminant Type Identification in Surface Electromyography Data*

Maurício C. Tosin, Leia B. Bagesteiro, and Alexandre Balbinot

Abstract — **This paper aims to present an innovative approach based on Reinforcement Learning (RL) concept to detect contaminants' type and minimize their effect on surface electromyography signal (sEMG). An agent-environment model was created based on the following elements: environment (muscle electrical activity), state (set of six features extracted from the signal), actions (application of filters/procedures to reduce the impact of each interference), and agent (controller, which will identify the type of contamination and take the appropriate action). The learning was conducted with Actor-Critic method. An average accuracy of 92.96% was achieved in an off-line experiment when detecting four contaminant types (electrocardiography (ECG) interference, movement artifact, power line interference, and additive white Gaussian noise).**

I. INTRODUCTION

The study of electromyography (EMG) has gained great attention from the scientific community because it consists of an easy way to understand how the force is muscle generated and how movements are performed from the physiological viewpoint [1]. Therefore, it is applied in several fields including clinical [2] and sports [3] medicine, physical therapy [4], and biomedical engineering [5], [6], to name a few. Many research projects are dedicated to Human-Machine-Interfaces (IHM) development such as myocontrolled prostheses [7], exoskeleton [8], and orthoses [9], among others. These studies are based on the achievement of patterns relating the myoelectric activity and user movement intention. Typically, machine learning algorithms are applied to create a predictive model for pattern recognition and motion decoding [10], [11].

High accuracy and efficient models are obtained when surface EMG signals contain the maximum amount of information. Thus, interference and noise level, which inevitably are present in electromyography recordings (due to instrumentation, environment, and physiological aspects) must be as low as possible. As a result, several research investigations proposed to detect, identify the type, and minimize the effect of contaminants in sEMG signals [12]–[14].

McCool and colleagues [12] presented a one class SVM based algorithm to recognize five contaminant types: electrocardiography (ECG) interference, motion artifact, powerline interference, saturation and additive white Gaussian noise. Their model training was performed by extracting seven features from the EMG recording.

Machado and colleagues [15] developed a hybrid algorithm based on Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM), which identified the contaminant font without the feature extraction stage. Whereas, Moura and Balbinot [13] applied a virtual sensor based on TVARMA (Autoregressive Moving Average) and TVK (Time-Varying Kalman) filter to a fault-tolerant classification system.

However, all previously cited methodologies require an off-line training stage and/or are not adaptable to environment variations. In this context, this work presents a new Reinforcement Learning (RL) based strategy to identify the type and attenuate the effect of contaminants in EMG records. RL is a methodology used in a wide variety of applications, going from gaming learning (e.g., backgammon) [16] to building automation (e.g., development of elevator traffic control system) [17]. Its application in sEMG pattern recognition yet scarce, is commonly used in proportional control of robot arm joint [18].

This work aims to propose a pre-processing strategy with the potential to identify contaminants' type and minimize their effect on sEMG signals. In addition, since the suggested algorithm is RL based, it is also adaptable to environmental changes and might be capable of performing online training. This paper is organized as follows: Section 2 details the database, contaminants considered in the study, feature extraction, and proposed algorithm. Section 3 shows obtained results. Discussion and conclusions are described in Section 4 and 5, respectively.

II. MATERIALS AND METHODS

RL is a learning strategy based on the interaction between an agent and an environment. For each action performed, the agent perceives a numeric reward and, in consequence the environment state may change. Thus, the goal is to get an action policy that maximizes the received long term reward using the agent acquired experience in its interaction with the environment.

An agent-environment model was created to adapt the RL concept to the sEMG signal processing conditions. In the context of this work, the environment is the sEMG recording,

* Research supported by National Counsel of Technological and Scientific Development (*CNPq* – Brazil).

M. C. Tosin and A. Balbinot are with the Electro-Electronic Instrumentation Laboratory of the Electrical Engineering Department, Federal University of Rio Grande do Sul, Porto Alegre, Rio Grande do Sul 90040-060 Brazil (mauricio.ctosin@gmail.com, abalbinot@gmail.com). L. B. Bagesteiro is with NeuroTech Lab, Kinesiology Department, San Francisco State University, San Francisco, CA, 94132, USA. (lbb1108@gmail.com).

the states are a set of six features extracted from the signal, the actions are filters/procedures applied to minimize each of the four contaminants considered, and the agent is the controller that will determine and apply the proper action.

A. Database

The publicly available repository NinaPro (Non-Invasive Adaptive Hand Prosthetics) was used to validate the proposed algorithm, more specifically, database 2 [19], exercise B, which counts with 17 different movements (8 basic finger motions and 9 basic wrist motions). This database is based on sEMG signals acquired from 40 intact participants through 12 electrode pairs. The first 10 volunteers records were analyzed in the present work.

B. Artificial Contamination

Four contaminants types were considered: electrocardiography (ECG) interference, motion artifact (MOA), powerline interference (PLI) and additive white Gaussian noise (WGN), which represents baseline noise. ECG recordings available at Physionet repository [20] were added to sEMG signals to generate the contaminated data according to procedure determined in [15]. PLI contamination was obtained adding a 50 Hz sinusoidal signal with random phase to the sEMG. MOA was generated using a simulation on a real acquisition, where each electrode was submitted to 10 taps with one second intervals to simulate an electrode displacement. A vector constituted by segments randomly selected from the motion artifact was used as the contamination to the sEMG signal. WGN was modeled according to [12]. A signal to noise ratio (SNR) of -30dB was employed in the contamination process. The SNR calculation was performed according to [15].

C. Feature Extraction

Six features were extracted in 500 ms windows without superposition. Four are frequency domain and were determined according to [21]: signal to motion artifact ratio (SMR), maximum drop in power density (DP), signal to noise ratio (SNR), and power spectrum deformation (DEF). The time domain metrics are signal to power line ratio (SPR) and signal to ECG ratio (SER), which were calculated as specified by [12].

D. Actions Definition

The actions consist of four filters/procedures to minimize the effect of each contaminant. A high-pass filter (20 Hz cutoff frequency) was applied for MOA and a low-pass filter (500 Hz cutoff frequency) for WGN interference. Both filters were 4[a] order Butterworth. A moving-average filter based on an algorithm proposed in [12] was implemented as ECG action. Lastly, the Least Squares Adaptive Algorithm presented in [22] was utilized to remove PLI.

E. Reward Function

For each action performed by the agent an individual score value function was determined to attribute a positive reward, if the correct filter/procedure was applied, and negative otherwise. The score value was determined according to the impact of the action on the value of specific features. The features considered for ECG, PLI and WGN scores determination were SER, SPR and SNR, respectively. The score associated with MOA was calculated using SMR, SER and SNR features.

Then, based on the score obtained for a specific action, a positive (+5), neutral (0), or negative (-10) reward was observed by the agent. If the score was above an upper limit, the reward was positive. If the score was between the upper and lower limits, the reward was neutral and for a score below the lower limit, the reward was negative. The upper and lower limits were determined according to score distributions quartiles values of each action.

F. Contaminant Identification Algorithm

Actor-Critic learning method was used to train the agent. This method consists of a parametrized state value function (critic) and a parametrized action policy (actor). The first determines an estimate for the long-term expected reward in a specific state, whereas the last one indicates a probability for the agent to take an action in a specific state. Thus, the training goal was to fit critic and actor parameters simultaneously to maximize total reward. For this, in each iteration both parameters were updated according to equations (1) and (2).

$$\boldsymbol{\theta_{t+1}} = \boldsymbol{\theta_t} + \alpha\delta_t \frac{\nabla_\theta \pi(A_t|S_t, \boldsymbol{\theta_t})}{\pi(A_t|S_t, \boldsymbol{\theta_t})} \tag{1}$$

$$\boldsymbol{w_{t+1}} = \boldsymbol{w_t} + \alpha\delta_t \nabla_w v(S_t, \boldsymbol{w_t}) \tag{2}$$

where, $\pi(A_t|S_t, \boldsymbol{\theta_t})$ is the action policy parametrized in θ, $\nabla_\theta \pi(A_t|S_t, \boldsymbol{\theta_t})$ denotes the gradient calculated as a function of θ, $v(s, \boldsymbol{w_t})$ is the state value function parameterized in w, $\nabla_w v(s, \boldsymbol{w_t})$ corresponds to the gradient taken in w, α is the learning factor, δ_t is the time difference error, given by Equation (3), and t is the time index.

$$\delta_t = R_{t+1} + \gamma \cdot v(S_{t+1}, \boldsymbol{w_t}) - v(S_t, \boldsymbol{w_t}) \tag{3}$$

where, R_{t+1} is the reward, S_t is the actual state, S_{t+1} is the next state and γ is a discount factor.

The critic model was defined as a linear function of the state. The action policy (actor) was implemented with a softmax function allowing probability values to be attributed for each action, as well as agent choices guide. It is defined by Equation (4).

$$\pi(a|s, \boldsymbol{\theta}) = \frac{\exp(h(s, a, \boldsymbol{\theta}))}{\sum_b \exp(h(s, b, \boldsymbol{\theta}))} \tag{4}$$

where $h(s, a, \boldsymbol{\theta})$ is a set of linear functions of the state (s) for each action (a) parametrized in $\boldsymbol{\theta}$. Therefore, h is composed of four equations, and $\boldsymbol{\theta}$ is its correspondent weight matrix.

Figure 1 shows the block diagram of the proposed ACRL (Actor-Critic Reinforcement Learning) based algorithm.

III. Results

The experiments were ran in MATLAB R2020b. The system model was developed with the Reinforcement Learning toolbox. The learning and discount factors were adjusted as 0.5 and 0.99, respectively. These parameters were selected after preliminary testing with the learning coefficient ranging from 0.005 to 1 and discount factor fixed at 0.99 [18].

The accuracies reported here correspond to the processing of each volunteer/channel pair. In each situation, the first half of the signal, previously contaminated with an intensity set at -30 dB, was used for training the agent and the remaining signal for the testing. Thus, all procedures were performed offline.

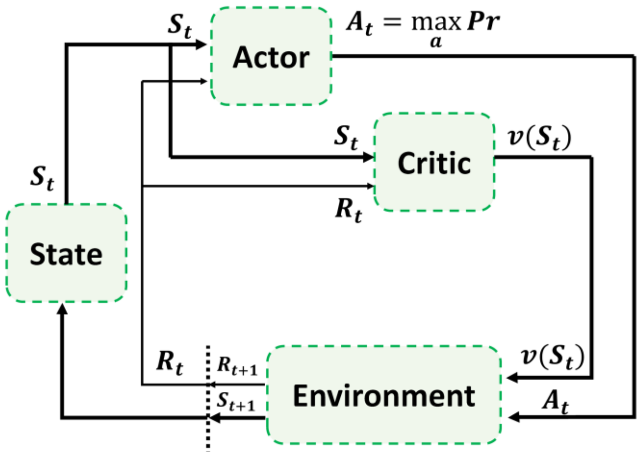

Figure 1. Schematic of proposed Actor-Critic algorithm.

A maximum number of episodes equal to 75 was determined as an algorithm stopping criterion, where one (1) episode indicates the processing of the entire data set for system training. This parameter was estimated empirically, guaranteeing, with a safety margin, that the training reaches the steady state. For each volunteer/channel pair, the algorithm was executed five times, totalizing 2400 results (10 volunteers x 4 contaminants x 12 channels x 5 repetitions).

Figure 2 shows the boxplot of all results obtained for each channel and contaminant. Each point in the boxplot corresponds to the average of the results obtained for all volunteers and repetitions. ANOVA test results showed no significant influence of volunteer in the algorithm assertiveness, while channel and contaminant were significant with 95% confidence level.

IV. DISCUSSION

A 92.96% global average accuracy was obtained applying the proposed method. The most easily identifiable was WGN contaminant, which presented the highest median and the lowest accuracy variability for all channels. It was followed by PLI and MOA, in that order. In contrast, the greatest oscillation in the results was shown by the ECG interference, with a 36.12% interquartile distance for channel #1. This is justified by the tendency of confusion with movement artifact as verified in some executions. It is noteworthy that the medians of all of four contaminants types were greater than 95% except for ECG and channel #1, indicating the effectiveness of the proposed recognition method.

In short, the results presented here are promising and comparable with other studies. The global average accuracy of the proposed algorithm (92.96%) is higher than the one obtained by Machado and collaborators [15], which registered 87.76% accuracy when identifying the same four types of contaminants plus a clean signal through a Recurring Neural Networks based methodology. It is important to notice that the current study used the signals of 38 volunteers from NinaPro Base 2 and processed them in 15 ms windows.

On the other hand, McCool and colleagues [12] reported an assertiveness rate of 100% in recognizing five sources of interference (ECG, motion artifact, powerline, white Gaussian noise and amplifier saturation) applying Support Vector Machine (SVM) algorithm. However, it is worth noting that the database used was smaller (113 records of 10 seconds), as compared to the one processed here (120 measurements of about 13.6 minutes). Besides, the level of contamination was different in these works, -20 dB as opposed to -30 dB. The SNR significantly influences the system performance, as demonstrated in previous works [12], [15], however the goal of these first tests was to verify the applicability of the algorithm in an extreme contamination situation (-30 dB). In the future, experiments under milder contaminations will be conducted to evaluate the algorithm effectiveness limit in identifying the type of contamination.

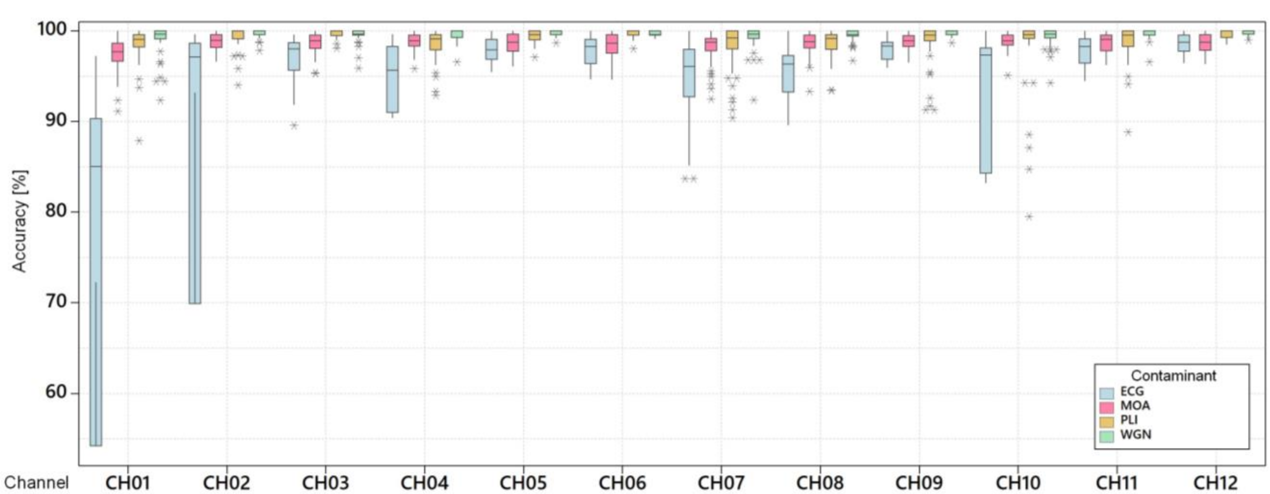

Figure 2. Boxplot of average accuracy obtained for each channel and contaminant.

V. CONCLUSION

The present study proposes a RL based algorithm to identify the source and minimize the effect of four common contaminants in sEMG signal. Some important aspects to encourage continuing in this line of research are: the potential for online learning brought with RL strategy, its adaptability to the environment changes, and the promising results achieved with the proposed algorithm pre-processing sEMG signal.

As a next stage, we will explore the online learning capability of the algorithm. For this, adaptations of the methodology will be performed and tested. Issues such as convergence time (data required to achieve convergence), computational complexity, robustness to unknown interferences, and the possibility of identifying more than one contaminant simultaneously present in the sEMG signal will also be investigated.

ACKNOWLEDGMENT

This study was partially supported by *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* – Brazil (*CAPES*) - Finance Code 001 and by *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (*CNPq* – Brazil).

REFERENCES

[1] C. J. De Luca, "The Use of Surface Electromyography in Biomechanics," *J. Appl. Biomech.*, vol. 13, no. 2, pp. 135–163, May 1997.

[2] A. Benazzouz and Z. E. H. Slimane, "Knee pathology diagnosis based on muscle activation intervals detection and the relationship between knee flexion and surface EMG," *Int. J. Med. Eng. Inform.*, vol. 13, no. 1, p. 14, 2021.

[3] S. di Fronso, C. Robazza, L. Bortoli, and M. Bertollo, "Performance Optimization in Sport: A Psychophysiological Approach," *Mot. Rev. Educ. Física*, vol. 23, no. 4, pp. 1–7, Nov. 2017.

[4] K. M. Shah, K. C. Madara, I. Diehl, M. Hyer, J. Schur, and P. W. McClure, "Shoulder muscle force and electromyography activity during make versus break tests," *Clin. Biomech.*, vol. 80, no. April, pp. 1–5, Dec. 2020.

[5] V. Cene, M. Tosin, J. Machado, and A. Balbinot, "Open Database for Accurate Upper-Limb Intent Detection Using Electromyography and Reliable Extreme Learning Machines," *Sensors*, vol. 19, no. 8, p. 1864, Apr. 2019.

[6] E. Trigili *et al.*, "Detection of movement onset using EMG signals for upper-limb exoskeletons in reaching tasks," *J. Neuroeng. Rehabil.*, vol. 16, no. 1, p. 45, Dec. 2019.

[7] A. Krasoulis, S. Vijayakumar, and K. Nazarpour, "Multi-Grip Classification-Based Prosthesis Control with Two EMG-IMU Sensors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 508–518, 2020.

[8] Z. Lu, K. Tong, X. Zhang, S. Li, and P. Zhou, "Myoelectric Pattern Recognition for Controlling a Robotic Hand: A Feasibility Study in Stroke," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 2, pp. 365–372, Feb. 2019.

[9] R. A. Bos, K. Nizamis, B. F. J. M. Koopman, J. L. Herder, M. Sartori, and D. H. Plettenburg, "A Case Study With Symbihand: An sEMG-Controlled Electrohydraulic Hand Orthosis for Individuals With Duchenne Muscular Dystrophy," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 258–266, Jan. 2020.

[10] V. H. Cene and A. Balbinot, "Enhancing the classification of hand movements through sEMG signal and non-iterative methods," *Health Technol. (Berl).*, vol. 9, no. 4, pp. 561–577, Aug. 2019.

[11] M. C. Tosin, V. H. Cene, and A. Balbinot, "Statistical feature and channel selection for upper limb classification using sEMG signal processing," *Res. Biomed. Eng.*, pp. 1–17, Aug. 2020.

[12] P. McCool, G. D. Fraser, A. D. C. Chan, L. Petropoulakis, and J. J. Soraghan, "Identification of contaminant type in surface electromyography (EMG) signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 774–783, 2014.

[13] K. de O. A. De Moura and A. Balbinot, "Virtual sensor of surface electromyography in a new extensive fault-tolerant classification system," *Sensors (Switzerland)*, vol. 18, no. 5, 2018.

[14] J. C. Machado, V. H. Cene, and A. Balbinot, "Recurrent Neural Network as Estimator for a Virtual sEMG Channel," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, no. 1, pp. 3620–3623.

[15] J. Machado, M. C. Tosin, L. B. Bagesteiro, and A. Balbinot, "Recurrent Neural Network for Contaminant Type Detector in Surface Electromyography Signals," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020, pp. 3759–3762.

[16] G. Tesauro, "Temporal difference learning and TD-Gammon," *Commun. ACM*, vol. 38, pp. 58–68, 1995.

[17] R. H. Crites and A. G. Barto, "Improving Elevator Performance Using Reinforcement Learning," *Adv. Neural Inf. Process. Syst.*, vol. 8, pp. 1017–1023, 1996.

[18] P. M. Pilarski, T. B. Dick, and R. S. Sutton, "Real-time prediction learning for the simultaneous actuation of multiple prosthetic joints," in *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*, 2013, pp. 1–8.

[19] M. Atzori *et al.*, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses.," *Sci. data*, vol. 1, pp. 1–13, 2014.

[20] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, Jun. 2000.

[21] C. Sinderby, L. Lindstrom, and A. E. Grassino, "Automatic assessment of electromyogram quality," *J. Appl. Physiol.*, vol. 79, no. 5, pp. 1803–1815, Nov. 1995.

[22] G. D. Fraser, A. D. C. Chan, J. R. Green, N. Abser, and D. MacIsaac, "CleanEMG - Power line interference estimation in sEMG using an adaptive least squares algorithm," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, no. 1, pp. 7941–7944.