Muscle Classification Via Hybrid CNN-LSTM Architecture from Surface EMG Signals

Esteban Velásquez and Jan Cornelis
Vrije Universiteit Brussel, Dept. ETRO
Pleinlaan 2, B-1050 Brussels, Belgium

Lubos Omelina, Bart Jansen
Vrije Universiteit Brussel, Dept. ETRO
Pleinlaan 2, B-1050 Brussels, Belgium
IMEC, Kapeldreef 75, B-3001 Leuven, Belgium

***Abstract*—Correct traceability of muscle identity within a predefined set of muscles in EMG studies is relevant in the periodic evaluation process of muscle training programs (for athletes), and in routine reviews for muscle rehabilitation. This article proposes hybrid deep learning CNN-LSTM models to classify the muscle directly from sEMG signals. These models allow for effective feature extraction and learning of short-term and long-term sequential dependencies. Two training setups are proposed: one using weight initialization provided from layer-wise unsupervised pretraining and another one using random initialization. Two validation scenarios are described to assess performance: testing on new contraction bursts from already-seen subjects in the training step (intrapersonal validation, useful in follow-up), and testing on a leave-one-out subject (interpersonal validation). Results indicate that the model can correctly classify different muscle groups in patients that already have been screened, but fails in distinguishing between symmetrical muscles.**

I. Introduction

Surface Electromyography (sEMG), is a technique that captures the electrical activity that is transferred to the surface of the skin, where non-invasive electrodes are located. This technique is currently employed for different types of applications, such as rehabilitation [1], advanced myoelectric control of prosthetic systems [2], gesture recognition [3], training program effectiveness, ergonomics [4], and movement analysis [5].

sEMG can be seen as an easier-to-use alternative for patients/users as it does not harm the soft tissue nor leave subsequent scarring, compared to other invasive EMG approaches. Also, this technique allows for unlimited reiteration of tests in the very same spot/muscle, making it ideal for the rehabilitation of neuromuscular disorders. Throughout an electromyography study, physicians and practitioners keep track of the muscles and exact locations where sEMG sensors are placed, either to assess the performance of the muscle or for periodic evaluation in the rehabilitation process. The traceability of the muscle involved is crucial and yet, it is currently done manually, making room for human errors and latency in the process.

In the present study, sEMG is employed to record electrical activity over different muscles of interest with the goal of capturing different muscle contraction bursts, from where a hybrid deep learning model can analyze patterns and infer the muscle on which the sensor is located (via classification). It's important to highlight that while muscles contract, the recorded sEMG contraction bursts are subject to anatomical and physiological factors that influence the signal such as muscle size, amount of motor unit potentials, demanded force, muscle fatigue, presence of abnormal behavior (fibrillation, fasciculation), existing conditions (myopathy or neuropathy), electrical interference, among others [6], [7]. These factors increase the complexity of muscle recognition tasks from sEMG signals, making deep learning models, such as the proposed hybrid CNN-LSTM, an appropriate candidate solution considering its capability of extracting multiple relevant features effectively, while also involving the time-series dependency inherent in sEMG recordings. To the best of our knowledge, identification of muscles directly from an sEMG signal has not been done and this paper introduces a framework to do so and demonstrates the accuracy in a feasibility study. One of the most direct exploitations of the results of this paper is the detection of errors in the measurement protocol of serial examinations, namely mixing up of the electrodes, or the omission of inputting the relation between the electrode number and the muscle on which it is placed.

II. Related Work

Concerning sEMG data processing, machine and deep learning techniques have been used to extract relevant features from the acquired signals for particular tasks, such as hand-gesture recognition and wrist kinematic estimation. In the study of Simao et al. [8], Feed-Forward Neural Network (FFNN), Recurrent Neural Network (RNN), Long Short-Term Memory network (LSTM) and Gated Recurrent Unit (GRU) are compared for the task of hand gesture classification from sEMG data obtained from UC2018 DualMyo and NinaPro DB5 datasets. These two datasets provide a series of activation bursts that are classified in the above mentioned hand gestures, which are ready to be used for supervised training in machine learning and deep learning models. The results obtained in this study indicate that FFNN, LSTM, RNN and GRU achieved similar accuracy for the previously mentioned datasets (around 95% for DualMyo and 91% for Ninapro). However, LSTM and GRU models used only a third of the parameters, compared to the other models, meaning smaller training and prediction times. Another study from Bao et al. [9] proposes a hybrid model, composed of CNN and LSTM, for EMG wrist kinematic estimation. For this purpose, sEMG data is collected from six healthy participants, where four wrist movements where performed. For the purpose of obtaining the kinematic estimation, a regression output model is built by means of integrating effective feature extraction

979-8-3503-3959-8/23/$31.00 ©2023 IEEE

provided by CNN, along with sequence regression by LSTM. The hybrid model is compared to CNN and LSTM alone, as well as with ML-based regression models such as Support Vector Regression (SVR) and Random Forest (RF). Results obtained in this study indicate CNN-LSTM's regressed output is smoother and considerably more similar to the actual ground truth than CNN alone, while also outperforming the previously mentioned models. Moreover, another study from Kumar et al. [10] also presents a hybrid deep learning approach for hand activities classification from sEMG signals. This study proposes the use of CNN, that are capable of extracting deep features from sEMG signals, along with Bidirectional LSTM, that extract bidirectional temporal information, for the purpose of correctly classifying different hand activities. Five different datasets were used for training and testing: NinaPro DB1, DB2, DB4, BioPatRec DB2 and UCI Gesture. The classification accuracy reached by the hybrid model on the previous datasets surpassed the 90% mark.

III. Methods

This study proposes the use of sEMG to record electrical activity over different muscles of interest with the goal of capturing a series of contraction bursts, from where a hybrid deep learning model is trained with the purpose of inferring on which muscle the sEMG sensor is located. To achieve this goal, the pipeline in Figure 1 is proposed.

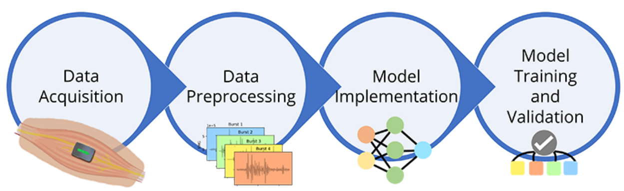

Fig. 1. Pipeline: Muscle Classification Via Hybrid Model

A. Data Acquisition

sEMG recordings of contraction bursts from different muscles are required to train and validate the model. For this purpose, Trigno Maize Sensor from the company Delsys is proposed and employed for this study [11]. Delsys electromyography instruments tackle the main challenges of this technology, such as noise signal artifact, contaminated signal from other muscles (crosstalk), and signal reliability and consistency. Moreover, their EMG instruments come with an easy to use software (Trigno Discover) for the data acquisition and storage. The selected sensor for this study was the Trigno Maize (dry electrode) grid sensor as it senses 16 different channels where the individual electrodes are organized in a 4×4 grid with approximately 5mm spacing. The sensor allows wireless data collection from all 16 electrodes in real-time with a sampling rate of 1000 Hz.

Concerning the experimental setup, 9 healthy volunteers participated in the study in Belgium: 5 men and 4 women ranging from 25 to 71 years old. None of the participants had any clinical history of neuromuscular disorders that could potentially affect the study. The steps proposed in the protocol as well as the possible risks related to the tests were given to the volunteers. The participants were allowed to stop/leave the session in case of any discomfort. The data were collected in a single session where 4 muscles were considered for the study: left and right biceps, as well as left and right tibialis anterior, whereby each recording involved a series of concentric and eccentric muscle contractions for each of these muscles. The 16 channel sensor was placed close to the muscle's centroid by considering specific anatomical landmarks and proportional distances from them (provided in a detailed protocol). The volunteers where asked to contract the muscle (concentric contraction) and hold the contraction for around 1 second, then release and go back to relaxed state (eccentric contraction) and repeat that every 5 seconds for a total duration of 10 minutes per muscle. This resulted into approximately 120 muscle contractions per muscle for each recording.

B. Data Preprocessing

The acquired sEMG signals come already prefiltered and normalized by the Delsys acquisition system: a high pass filter (2-pole Butterworth) and low pass filter (8-pole Butterworth). Some other prefiltering considerations applied to the sEMG signals can be seen in the Delsys sEMG Detection and Recording manual [12]. Throughout the duration of each of the sEMG recordings, two main behaviors can be identified in the signal: the rest period, in which the muscle is in a relaxed state and the electrical activity does not deviate too much from the baseline; and the contraction burst period, where few or multiple motor units activate different muscle fibers, causing an increment in the signal amplitude and frequency. In order to identify the contraction bursts in the sEMG recordings, the beginning and end of each phase of each muscle activation is determined by a single threshold algorithm that makes use of the Teager Kaiser Energy Operator (TKEO), which suppresses the noise and makes the signal larger; thus, making the bursts more easily identifiable. The threshold was chosen empirically as a single value for the whole dataset. Figure 2 shows the detected activation signals obtained from the proposed algorithm. Since contraction bursts' duration can vary substantially between subjects and even for the same subject throughout the test, only 500 milliseconds of the signal are considered (prioritizing the onset of the contraction).

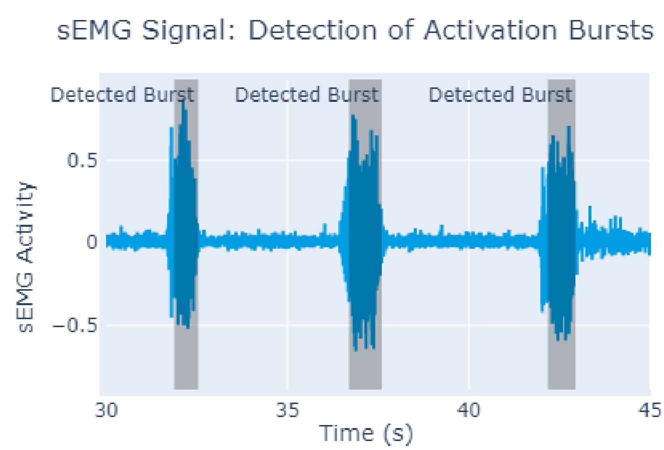

Fig. 2. Recorded sEMG signal from a single channel with highlighted sequences of muscle contraction bursts as detected by the detection algorithm.

The 16 channels offered by the sEMG sensor in the two-dimensional grid configuration are relatively close to each other (6 mm). Thus, the electrical activity in an activation burst

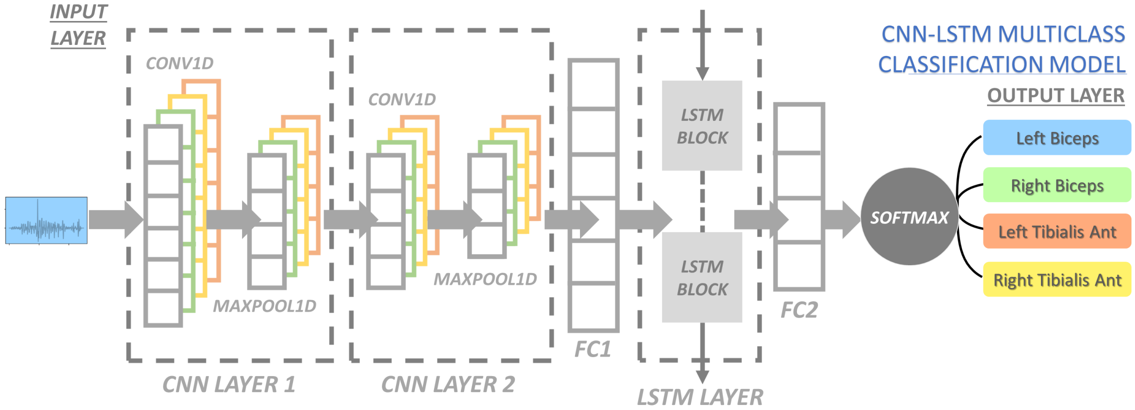

Fig. 3. Hybrid CNN-LSTM Model Architecture

does not deviate dramatically between these channels but rather follows a similar pattern. In this sense, the 16 signals identified in each burst were used independently (as separate samples), as a sort of data augmentation strategy. From the total 4361 contraction bursts obtained from the experiments with the participants, a total of 69776 data samples are obtained after splitting bursts into the 16 channels. These samples are used for the training and validation of the hybrid deep learning model.

C. Model Implementation

As can be concluded from the state-of-the-art analysis, sEMG signal processing is nowadays being approached by hybrid deep learning techniques with better results compared to more traditional machine learning techniques. The models with higher performance normally involved the use of CNN and LSTM for tasks such as gesture recognition and wrist kinematics estimation. Based on this, a hybrid CNN-LSTM is proposed in the current paper as the architecture to address the muscle classification task. The proposed model is composed of CNN, LSTM and FC layers, taking advantage of CNN layers as feature extractors and the LSTM layer to maintain long-term dependencies that are inherent in sEMG signals. Figure 3 represents the architecture employed. The input tensor goes through two CNN layers followed by a FC layer, similar to one of the architectures presented by Xiang Chen et al.[13], which then goes to one LSTM layer. Notice that LSTM layers are prone to overfit more easily than CNN layers [9], and for this reason only one layer of LSTM is defined for this study. Finally the output of the LSTM goes through a second FC layer and then a last FC layer with a softmax function providing the output tensor, which corresponds to the probability of the four muscle classes. Take into account that some of the hyperparameters are determined via empirical manual tuning.

The dataset distribution for the training and validation of the model contemplates two validation scenarios: In the first one, contraction bursts from eight subjects are used in the training of the model (80% of samples) and later the model is validated (20% of samples) on new contraction bursts from previously seen participants in the training step (intrapersonal validation). In the second validation scenario, the ninth subject (leave-one-out) is used to assess the model's ability to correctly classify new contraction bursts from unseen participants in the training step (interpersonal validation); this second validation scenario was repeated with different leave-one-out subjects, leading to similar results in all of them. Note that the collected dataset was obtained from a single session.

D. Model Training and Validation

Due to lack of publicly available pretrained models processing the sEMG signal, it was decided to use autoencoders (AE) to obtain optimal initialization weights for part of the designed network. A Greedy layer-by-layer unsupervised pre-training strategy [14] is proposed to obtain the weight initialization of the two CNN layers. This method is known to provide great generalization properties and to overcome problems of local optimization during the training phase. In this vein, AE architectures are built to implement a reconstruction task for each of the CNN layers, meaning that each of these layers is isolated and used in the encoder part of its own specific AE for recontsruction. An encoder-decoder architecture is thus put in place, where the encoder is driven by the CNN layer, while the decoder is attached in form of a deconvolution operation that allows reconstructing the input tensor from the latent representation. The parameters learned by the encoder part from both CNN layers are posteriorly transferred and used in the final model as illustrated in Figure 4.

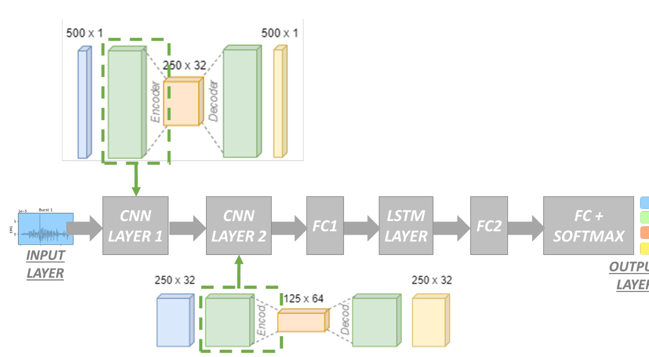

Fig. 4. Parameter Transfer from Pretraining to the Final Model (Weight Initialization)

Optimal weights are obtained after training the AE for the reconstruction task and these are used to initialize part of the designed model

Multiclass Classification Task

Intrapersonal Validation: Train Setup 1 (Pretraining weights)

Confusion Matrix (Ground Truth \ Prediction)	L Biceps	R Biceps	L Tibialis	R Tibialis
L Biceps	2872	65	6	1
R Biceps	145	2791	5	3
L Tibialis	2	1	3041	172
R Tibialis	19	1	227	3113
			Accuracy	95%

Intrapersonal Validation: Train Setup 2 (Random weights)

Confusion Matrix (Ground Truth \ Prediction)	L Biceps	R Biceps	L Tibialis	R Tibialis
L Biceps	2666	255	20	3
R Biceps	128	2809	5	2
L Tibialis	1	1	2958	256
R Tibialis	16	0	282	3062
			Accuracy	92%

Interpersonal Validation: Train Setup 1 (Pretraining weights)

Confusion Matrix (Ground Truth \ Prediction)	L Biceps	R Biceps	L Tibialis	R Tibialis
L Biceps	303	498	21	186
R Biceps	696	417	6	273
L Tibialis	55	11	924	930
R Tibialis	16	5	563	1336
			Accuracy	48%

Interpersonal Validation: Train Setup 2 (Random weights)

Confusion Matrix (Ground Truth \ Prediction)	L Biceps	R Biceps	L Tibialis	R Tibialis
L Biceps	487	1332	21	80
R Biceps	608	1238	4	54
L Tibialis	35	82	1467	336
R Tibialis	576	121	733	490
			Accuracy	48%

a)

Binary Classification Task Biceps vs Tibialis Ant.

Intrapersonal Validation with Pretraining weights

Confusion Matrix (GT \ Prediction)	Biceps	Tibialis
Biceps	5876	12
Tibialis	3	6573
	Accuracy	99%

Interpersonal Validation with Pretraining weights

Confusion Matrix (GT \ Prediction)	Biceps	Tibialis
Biceps	3684	140
Tibialis	699	3141
	Accuracy	89%

Binary Classification Task Left vs Right Muscles

Intrapersonal Validation with Pretraining weights

Confusion Matrix (GT \ Prediction)	Left	Right
Left	5746	414
Right	468	5836
	Accuracy	93%

Interpersonal Validation with Pretraining weights

Confusion Matrix (GT \ Prediction)	Left	Right
Left	2464	1376
Right	2469	1355
	Accuracy	50%

b)

Fig. 5. Confusion Matrices: Multiclass Classification (a) Binary Classification (b).

IV. Results

Two training setups are presented here: A first one using the weight initialization provided from the layer-wise unsupervised pretraining step and the second one that only uses random weight initialization. After training the model for 200 epochs, the model was able to achieve 95.0% (setup 1) and 92.56% (setup 2) classification accuracy in the validation dataset (intrapersonal scenario), where new contraction bursts from already seen subjects where given to the model for the validation purposes. Concerning the second proposed validation scenario (interpersonal scenario or leave-one-out), the model achieved 48% accuracy (in both setups) for muscle classification of contraction bursts from an excluded subject in the training. The evolution of the loss, as well as the validation accuracy over the epochs is presented in Figure 6. Something interesting that can be seen from Figure 5 (a) is that when there is a misclassification in the interpersonal validation, it is often between symmetrical muscles, meaning for instance that left biceps is more likely to be misclassified as right biceps and viceversa (also for left and right tibialis anterior). This behaviour ends up heavily penalizing the overall accuracy of the model in this second validation scenario and it is confirmed by performing two additional tests in which the multiclass classification problem is reduced to a binary classification problem (biceps vs tibialis anterior, and left muscles vs right muscles). Confusion matrices from these two additional tests are included in Figure 5 (b). Further follow-up recordings would be essential in future studies to determine the intersession performance of the model.

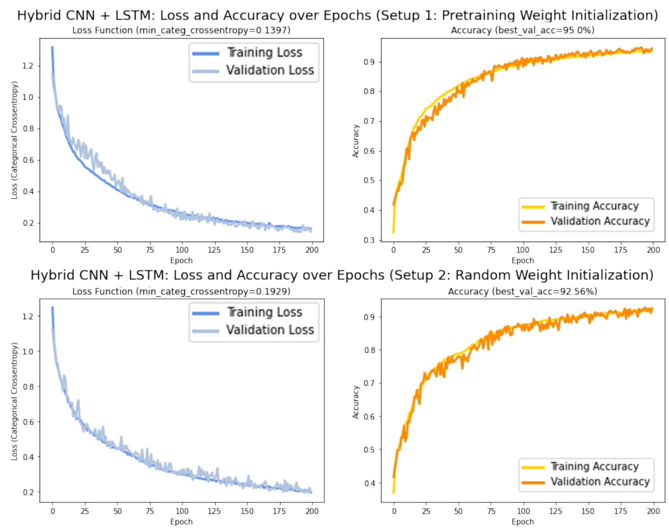

Fig. 6. Training and Intrapersonal Validation: Loss and Accuracy over Epochs for the two training setups

V. Conclusion

The goal of this article is to investigate the possibility of classifying the underlying muscles directly from the sEMG signal. This is motivated by the fact that the correct traceability of the implicated muscles in EMG studies is relevant in the periodic evaluation process that is carried out for instance in muscle training program effectiveness for athletes, as well as in routine reviews for muscle rehabilitation.

The proposed hybrid CNN-LSTM model was validated in two scenarios: in the first one, the model is tested on new contraction bursts from already seen subjects in the training step, where at least 9 out of 10 times the muscle category was accurately inferred (in both training setups). On the other hand, in the leave-one-out validation scenario, the model could barely classify correctly half of the times, as it seems incapable of distinguishing between symmetrical muscles. When constraining the task to a binary classification problem of biceps vs tibialis anterior, the model is able to classify these non-symmetrical muscles almost nine out of ten times (89% accuracy). On the other hand, when switching the binary classification task to left muscles vs right muscles, the model only achieves 50% classification accuracy. From our feasibility study we conclude that muscle category can be estimated, but not the lateral side as most of the muscles are mirrored.

Acknowledgment

The authors of this study thank the Vrije Universiteit de Brussel (VUB) for providing the resources and support needed to complete the research. The authors would also like to extend their gratitude to the volunteers who participated in the data acquisition step and to the VUB ETRO department staff for their significant contributions to this research.

REFERENCES

[1] C. L. Toledo-Peral, G. Vega-Martínez, J. A. Mercado-Gutiérrez, G. Rodríguez-Reyes, A. Vera-Hernández, L. Leija-Salas, and J. Gutiérrez-Martínez, "Virtual/augmented reality for rehabilitation applications using electromyography as control/biofeedback: Systematic literature review," *Electronics*, vol. 11, no. 14, p. 2271, 2022.

[2] N. Parajuli, N. Sreenivasan, P. Bifulco, M. Cesarelli, S. Savino, V. Niola, D. Esposito, T. J. Hamilton, G. R. Naik, U. Gunawardana *et al.*, "Real-time emg based pattern recognition control for hand prostheses: A review on existing methods, challenges and future implementation," *Sensors*, vol. 19, no. 20, p. 4596, 2019.

[3] 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.

[4] F. Barroso, D. R. Bueno, J. Á. Gallego, P. Jaramillo, and A. Kilicarslan, "Surface emg in neurorehabilitation and ergonomics: State of the art and future perspectives," *Emerging Therapies in Neurorehabilitation*, pp. 267–284, 2014.

[5] N. Massó, F. Rey, D. Romero, G. Gual, L. Tutusaus, and A. Germán Romero, "Surface electromyography applications," *Apunts: Medicina de l'esport; Vol.: 44 Núm.: 166*, vol. 45, 01 2010.

[6] J. Kimura, "Electrodiagnosis in diseases of nerve and muscle: principles and practice," 2013.

[7] S. Rampichini, T. M. Vieira, P. Castiglioni, and G. Merati, "Complexity analysis of surface electromyography for assessing the myoelectric manifestation of muscle fatigue: A review," *Entropy*, vol. 22, no. 5, p. 529, 2020.

[8] M. Simão, P. Neto, and O. Gibaru, "Emg-based online classification of gestures with recurrent neural networks," *Pattern Recognition Letters*, vol. 128, pp. 45–51, 2019.

[9] T. Bao, S. A. R. Zaidi, S. Xie, P. Yang, and Z.-Q. Zhang, "A cnn-lstm hybrid model for wrist kinematics estimation using surface electromyography," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, 2020.

[10] N. K. Karnam, S. R. Dubey, A. C. Turlapaty, and B. Gokaraju, "Emghandnet: A hybrid cnn and bi-lstm architecture for hand activity classification using surface emg signals," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 1, pp. 325–340, 2022.

[11] Delsys, "Delsys trigno maize, base station and trigno discover software," 2022, [Online; accessed June, 2022]. [Online]. Available: https://delsys.com/trigno-maize/

[12] Delsys, "Delsys surface electromyography: Detection and recording," 2002, [Online; accessed June, 2022]. [Online]. Available: https://www.delsys.com/downloads/TUTORIAL/semg-detection-and-recording.pdf

[13] X. Chen, Y. Li, R. Hu, X. Zhang, and X. Chen, "Hand gesture recognition based on surface electromyography using convolutional neural network with transfer learning method," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 4, pp. 1292–1304, 2020.

[14] M. Ali, D.-H. Son, S.-H. Kang, and S.-R. Nam, "An accurate ct saturation classification using a deep learning approach based on unsupervised feature extraction and supervised fine-tuning strategy," *Energies*, vol. 10, no. 11, p. 1830, 2017.