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

Estimation of Joint Angle From sEMG and Inertial Measurements Based on Deep Learning Approach

Alfredo Lobaina Delgado[1], Adson F. da Rocha[2], Alexander Suárez León[1], Andrés Ruiz-Olaya[3], Klaus Ribeiro Montero[2] and Alberto López Delis[4]

Abstract— **Continuous kinematics estimation from surface electromyography (sEMG) allows more natural and intuitive human-machine collaboration. Recent research has suggested the use of multimodal inputs (sEMG signals and inertial measurements) to improve estimation performance. This work focused on assessing the use of angular velocity in combination with myoelectric signals to simultaneously and continuously predict 12 joint angles in the hand. Estimation performance was evaluated for five functional and grasping movements in 20 subjects. The proposed method is based on convolutional and recurrent neural networks using transfer learning (TL). A novel aspect was the use of a pretrained deep network model from basic joint hand movements to learn new patterns present in functional motions. A comparison was carried out with the traditional method based solely on sEMG. Although the performance of the algorithm slightly improved with the use of the multimodal combination, both strategies had similar behavior. The results indicated a significant improvement for a single task: opening a bottle with a tripod grasp.**

Index Terms— **Angle estimation, angular velocity, deep learning, inertial measurements, recurrent and convolutional neural networks (RCNN), surface electromyography (sEMG), transfer learning (TL)**

I. INTRODUCTION

Rehabilitation and assistive technologies are often aimed at the restoration and functional compensation of body structures after injuries or neuromuscular disability. These technologies are based on human-robot collaboration [1]. The developers have focused on improving the performance of upper limb exoskeletons and prostheses controlled by surface electromyography (sEMG). The sEMG signals are still used due to their easy application, noninvasiveness, neural information, and wide use for human-machine interfacing in clinical practice [2]–[6].

The human motor system performs movements with multiple degrees of freedom (DoF) simultaneously during the execution of kinetic chain movements. A suitable strategy based on sEMG to address this challenge is continuous and simultaneous motion estimation [3], [4]. This paradigm provides a more natural and intuitive way for robotic devices that mimic body movements [4].

Despite the usefulness of sEMG signals as input sources, the estimation algorithm performance is influenced by several factors, such as the change of position of the extremities and their orientation [5], [7]. One of the current solutions for the limb position effect is the multimodal combination using inertial sensors as an additional source of information [5], [7]. sEMG sensors, in addition to Inertial Measurements Units (IMUs) have been used for predicting joint angles [8]. IMUs usually comprise several sensors, such as gyroscopes, accelerometers, and magnetometers. These sensors provide information about the movement after it has been executed, while sEMG signals allow the prediction of motor intent [8]. Recent research work (e.g., [9]–[13]) has suggested the hypothesis that models based on multisensory modality outperforms conventional method based exclusively on sEMG.

Most estimation methods are developed using artificial neural networks (ANN) [2]–[4], [13]–[16]. Even so, few research works have used multimodal combinations. Nevertheless, the precision of these algorithms in estimating motion kinematics is still inadequate to be implemented in practice [17], [18]. Currently, deep neural network-based techniques have been published: [19], [20]. These *deep learning* methods allow automatic feature extraction [21]. In addition, they allow working with large amounts of data favoring the use of other sources of information.

Some of the deep learning algorithms in the myoelectric control paradigms are the following: convolutional networks (CNNs), recurrent networks (RNNs), as well as the fusion of both: RCNN [5], [6], [21]. One of the limitations of CNNs is that they can only process data in a single window, thus losing the ability to express time-dependent features [18]. RNNs have the advantage of being able to perform sequential processing, which is ideal for myoelectric and inertial signals due to their temporal nature [21]. RCNN combines the advantages of RNNs and CNNs, optimizing the deep architecture for better performance [18]. This model has been selected for the estimation carried out in the present study.

This work describes a method based on deep neural networks and the multimodal combination: sEMG + IMU. The aim is to evaluate the effect of the inclusion of inertial measurements as another source of information to enhance the performance of continuous estimation algorithms compared to the traditional use of myoelectric signals.

[1]Alfredo Lobaina Delgado and Alexander Suárez León are with the Department of Biomedical, Faculty of Telecommunications, Biomedical and Informatics Engineering, Universidad de Oriente, Santiago de Cuba, Cuba, alobainad@gmail.com

[2] Adson F. da Rocha and Klaus Ribeiro Montero are with the Department of Electrical Engineering, University of Brasilia, Brasilia, Brazil, adsonr@gmail.com

[3] Andrés Ruiz-Olaya is with Antonio Nariño University, Colombia

[4]Alberto López Delis is with the Center of Medical Biophysics, Universidad de Oriente, Santiago de Cuba, Cuba, lopez.delis69@gmail.com

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

II. MATERIALS AND METHODS

A. Ninapro database

The dataset used was obtained from the seventh Ninapro database (DB7), collected by Krasoulis *et al.* and described in [7]. The data used for training and testing comprised sEMG signals and kinematic information from twenty right-handed and able-bodied subjects. Myoelectric signals and IMU data were collected using 12 sEMG electrodes, and each one incorporated a tri-axial accelerometer, gyroscope, and magnetometer measuring acceleration, angular velocity, and magnetic field, respectively. Raw gyroscopes values $[deg/s]$ were used as complementary input. All data were synchronized via linear interpolation to the highest sampling frequency (i.e., 2000Hz). Hand angles were recorded by using a dataglove (Cyberglove II, 18-degree of freedom (DOF)). Twelve joints were selected for estimation: five proximal interphalangeal (PIP), five metacarpophalangeal (MCP), the trapeziometacarpal (TMC) and the wrist yaw (WY) joint angles. These joints are used in most myoelectric commercial hand prostheses. Five continuous classes corresponding to grasping and functional movements were predicted, including rest periods: (**A**)- *index finger extension*, (**B**)- *ring*, (**C**)- *power*, (**D**)- *parallel extension*, and (**E**)- *open a bottle with a tripod grasp* (i.e., **21**, **23**, **27**, **35**, and **38** respectively, from exercise 2 included in DB7). The basic finger and wrist movements were used in the pre-training for transfer learning.

B. Preprocessing

In order to remove the movement artifacts, all sEMG channels were bandpass-filtered (4th order Butterworth with 20–500Hz cutoff frequencies). The signals were rectified and then low-pass filtered by a 2nd order Butterworth filter with cut-off frequency to 6 Hz to obtain the linear envelope. This last filter was also used for angular velocity measurements processing.

C. Transfer learning (TL)

The proposed model in this work is based on recurrent and convolutional neural networks, which require training with a significant number of samples. Although the database is extensive, each class has only six movement repetitions. In order to consider the processing window size of each sequence, the training sets are made small compared to the total number of weights to be adjusted by neural networks.

Due to these limitations, transfer learning was performed. The hypothesis is that a pre-trained deep network model from basic joint movements can be used as a starting point to learn new patterns present in functional motions. Then, the learned features are transferred to a new motor task using a smaller number of training observations. TL can be applied from data obtained across multiple subjects [6], but there are individual differences in the subjects' anatomy. Therefore, using data from basic movements can also improve the performance of movement estimation with more significant variation.

The pre-trained networks (PTNs) were obtained using the signals acquired during isometric and isotonic hand contractions exercises and basic movements of the wrist (i.e., classes 1 to 17, whole exercise 1).

1) PTN architecture design: The proposed network (see Table I) includes a convolution, batch normalization and ReLU blocks, one Long Short-Term Memory (LSTM) recurrent layer that outputs the last element of a sequence, and one dense layer. In order to perform the convolutional operations on each time step independently, the structure includes a sequence folding layer before the convolutional layers. Since the LSTM layer expects sequences of vectors as inputs, the output of the convolutional layers is reshaped to sequences of feature vectors by inserting a sequence unfolding layer and a flatten layer between the convolutional layer and the LSTM one.

TABLE I

RCNN PROPOSED MODEL

No	Layers	Parameters
1	Sequence Input:	2D Window: channels x samples
2	Sequence Folding	
3	Convolution 2D:	Kernel size: 3, Number of filters: 64, Stride: 1
4	Batch Normalization	
5	ReLU	
6	Sequence Unfolding	
7	Flatten	
8	LSTM:	Hidden states: 128
9	Dropout:	Probability: 0.5
10	Fully Connected:	Output size: 12
11	Regression:	Loss Function: *Mean Square Error*

2) Training settings: Both models (PTN and the proposed network) were adjusted using stochastic gradient descent with momentum. The initial learning rate was set to 0.01. The gradient threshold was set to 1 because it prevents the gradients from exploding. In order to avoid over-fitting, batch normalization, dropout, and early stopping were used.

3) Procedure: Although there are several ways of carrying out TL, the one used in this work is the most common approach, in which it keeps the features from the earliest layers of the pre-trained network (the transferred layer weights). Furthermore, the final layers are replaced with new layers adapted to the new data. Then, the learning rate factors were set to 10 for the new dense layer. In this way, the layer learns faster than the other transferred layers. The replaced layers were the last fully connected layer and the final regression layers of the PTN. Fig.2 shows the structure of the deep neural network.

The data were normalized to have zero mean and unit variance. The test and validation data were also standardized using the same training parameters. The input sequence is a matrix constructed on a segment of the sEMG and IMU channels. The sliding window method was used to define the time window. In this case, the length of the shifting window was set to 50 ms (100 samples) and the increment

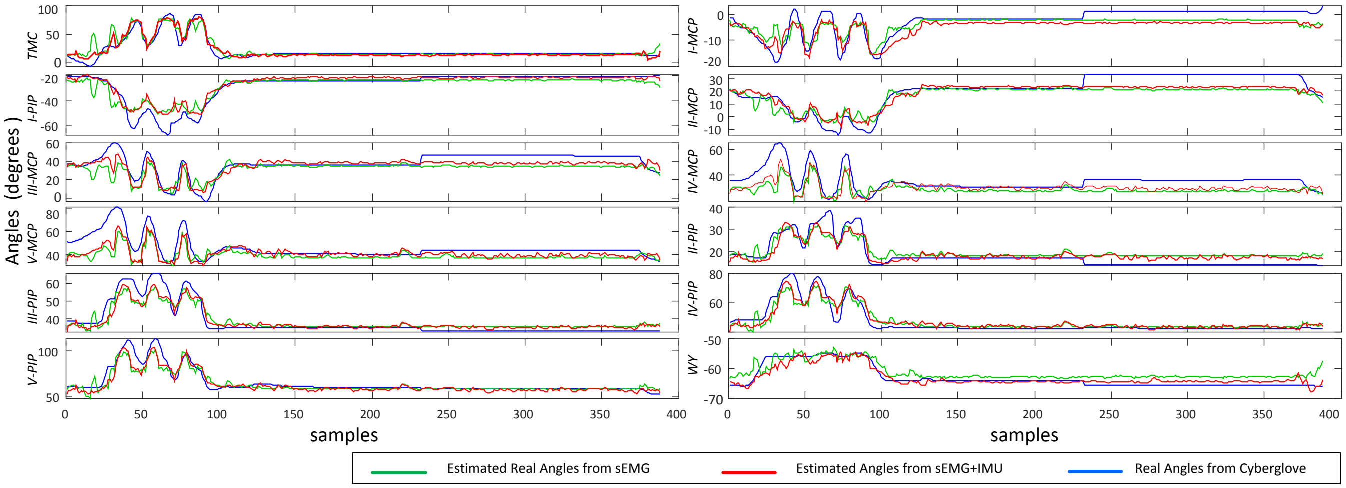

Fig. 1. Estimated and real joint angles from multimodal combination and traditional method. The data corresponds to Subject 10 during last repetition of the movement **E**.

to 30 ms (60 samples). Four repetitions were used to create the training set, one as the validation set and the remaining repetition as the test set.

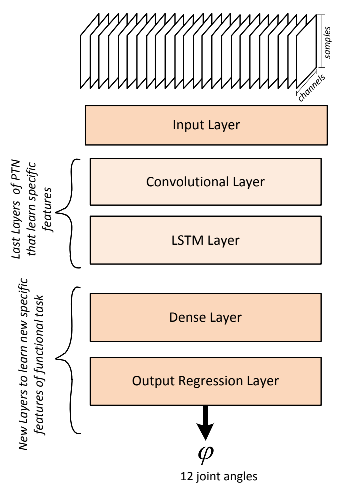

Fig. 2. Transfer learning to retrain a RCNN

III. RESULTS AND DISCUSSION

The performance evaluation of the proposed model was based on widely used regression metrics. The following metrics were used: correlation coefficient (CC), global coefficient of determination (R^2), and normalized root mean square error (NRMSE). CC can measure the similarity between signal shapes, while global R^2 evaluates the performance from all the joints involved. NRMSE was calculated from residual prediction error between the estimated and measured angles. NRMSE is used to standardize the RMSE to the difference between maximum and minimum real data. The values of global R^2 and CC closer to 1 indicate good accuracy, and smaller values of NRMSE indicated that the model has a better performance.

Fig.1 shows a qualitative comparison between joint angles estimated from sEMG and the multimodal combination versus measured angles. In this example, the mean values of the global R^2 and CC estimates from sEMG (i.e., 0.62 and 0.85, respectively) are less than the estimate from sEMG+IMU (i.e., 0.75 and 0.89). The opposite happens with the NRMSE mean value across joints, which is higher for the model from sEMG (0.14) than for the combination (0.12). The main error in the estimation from the conventional method occurs at the beginning of the activation period, but, still, in both information source comparisons, the model convergence at the end of the rest period is almost the same and reasonably close to the actual measurements. In most cases, the midline and amplitude range from predicted angles match the recorded ones.

An analysis of variance (ANOVA) was performed in order to test whether the average values of each performance index were significantly different for both types of inputs, i.e. sEMG-only and sEMG+IMU. The average was calculated from the values obtained for each joint across all subjects. The Kolmogorov-Smirnov test was applied, and it does not reject the hypothesis that the data come from a normal distribution at the 1% significance level. In the ANOVA test, the CC with sEMG+IMU (0.70 ± 0.12) is higher than sEMG (0.64 ± 0.11) for movement **E**, and the means both are significantly different (p=0.000048). This kind of movement presents more variation in joint angle shapes than the others. A median analysis also corroborates these results. Fig.3 shows that the notches in the box plot overlap in movements **A** to **D**. Thus, the results suggest that the medians differ in movement **E**. However, for the remaining movements, there were no significant differences.

Table II shows the statistics of NRMSE and global R^2 across all subjects and joints. Regarding the sample means, there were no differences for all movements according to the ANOVA test. Wang *et al.*, in [22], obtained similar NRMSE values from sEMG using an RNN based on LSTM layers.

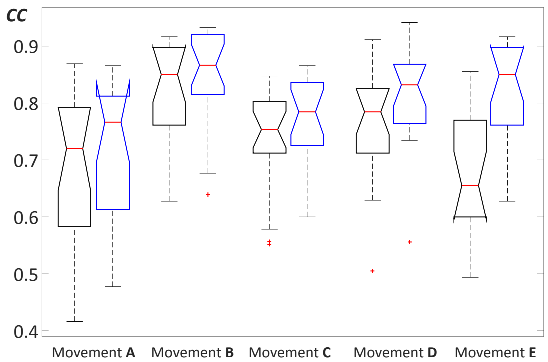

Fig. 3. Differences between the medians of the CC values across all subjects for both input sets (sEMG and sEMG+IMU). The blue boxes correspond to the multimodal combination while the black boxes correspond to sEMG. The central mark in the boxes represents the median; the edges of the boxes are the 25th and 75th percentiles.

Concerning estimation accuracies from global R^2, the results were low. These results might be because this performance index not only allows quantifying the continuous movement, but also the simultaneous movement from all joints and their DoFs. So for functional movements, the estimation performance is still a challenge. Even so, there were subjects with performance between 0.75 and 0.85. For the multimodal combination, more subjects exceeded the value of 0.75 than for the conventional method.

In all movements and joints, the multimodal combination improves the angle estimation in comparison with traditional methods. Even so, the statistics show similar results for both comparison criteria.

TABLE II
NRMSE AND GLOBAL R^2, (MEAN $\pm$ STANDARD DEVIATION)

Classes	NRMSE		global R^2	
	sEMG	*sEMG+IMU*	*sEMG*	*sEMG+IMU*
A	**0.191**$\pm$0.037	**0.185**$\pm$0.039	**0.349**$\pm$0.270	**0.391**$\pm$0.216
B	**0.199**$\pm$0.041	**0.188**$\pm$0.040	**0.624**$\pm$0.182	**0.664**$\pm$0.175
C	**0.223**$\pm$0.041	**0.217**$\pm$0.049	**0.539**$\pm$0.217	**0.553**$\pm$0.257
D	**0.212**$\pm$0.057	**0.202**$\pm$0.059	**0.464**$\pm$0.234	**0.520**$\pm$0.256
E	**0.218**$\pm$0.046	**0.214**$\pm$0.048	**0.636**$\pm$0.191	**0.663**$\pm$0.190

REFERENCES

[1] L. Bi, C. Guan, *et al.*, "A review on EMG-based motor intention prediction of continuous human upper limb motion for human-robot collaboration," *Biomedical Signal Processing and Control*, vol. 51, pp. 113–127, 2019.

[2] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, pp. 797–809, July 2014.

[3] J. Liu, S. Kang, D. Xu, Y. Ren, S. L, and L.-Q. Zhang, "EMG-based continuous and simultaneous estimation of arm kinematics in able-bodied individuals and stroke survivors," *Frontiers in Neuroscience*, vol. 11, August 2017.

[4] Q. Zhang, R. Liu, W. Chen, and C. Xiong, "Simultaneous and continuous estimation of shoulder and elbow kinematics from surface EMG signals," *Frontiers in Neuroscience*, vol. 11, May 2017.

[5] E. Campbell, A. Phinyomark, and E. Scheme, "Current trends and confounding factors in myoelectric control: Limb position and contraction intensity," *Sensors*, 2020.

[6] A. Phinyomark, E. Campbell, and E. Scheme, "Surface electromyography (EMG) signal processing, classification, and practical considerations," in *Biomedical Signal Processing. Advances in Theory, Algorithms and Applications* (G. Naik, ed.), pp. 3–29, Springer, 2020.

[7] A. Krasoulis, I. Kyranou, M. S. Erden, K. Nazarpour, and S. Vijayakumar, "Improved prosthetic hand control with concurrent use of myoelectric and inertial measurements," *Journal of NeuroEngineering and Rehabilitation*, vol. 14, p. 71, Jul 2017.

[8] K. Zhang, C. W. de Silva, and C. Fu, "Sensor fusion for predictive control of human-prosthesis-environment dynamics in assistive walking: A survey," *arXiv*, 2019.

[9] A. Krasoulis, S. Vijayakumar, and K. Nazarpour, "Evaluation of regression methods for the continuous decoding of finger movement from surface EMG and accelerometry," in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 631–634, April 2015.

[10] D. Blana, T. Kyriacou, J. M. Lambrecht, and E. K. Chadwick, "Feasibility of using combined EMG and kinematic signals for prosthesis control: A simulation study using a virtual reality environment," *Journal of Electromyography and Kinesiology*, vol. 29, pp. 21–27, 2016.

[11] A. López D., C. J. Miosso, J. a. L. A. Carvalho, A. F. da Rocha, and G. A. Borges, "Continuous estimation prediction of knee joint angles using fusion of electromyographic and inertial sensors for active transfemoral leg protheses," *World Scientific*, vol. 10, no. 2, 2018.

[12] X. Zhang, X. Tang, X. Zhu, X. Gao, X. Chen, and X. Chen, "A regression- based framework for quantitative assessment of muscle spasticity using combined EMG and inertial data from wearable sensors," *Frontiers in Neuroscience*, vol. 13, 2019.

[13] A. Lobaina Delgado, C. Quinayás, A. Ruiz, A. F. da Rocha, and A. López Delis, "Hand angle estimation based on sEMG and inertial sensor fusion," in *Smart Technologies, Systems and Applications*, 2020.

[14] J. M. Hahne, F. Biemann, N. Jiang, H. Rehbaum, D. Farina, F. C. Meinecke, K. R. Mller, and L. C. Parra, "Linear and nonlinear regression techniques for simultaneous and proportional myoelectric control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, pp. 269–279, 2014.

[15] R. Raj and K. Sivanandan, "Comparative study on estimation of elbow kinematics based on EMG time domain parameters using neural network and anfis narx model," *Journal of Intelligent & Fuzzy Systems*, 2016.

[16] I. Vujaklija, V. Shalchyan, E. N. Kamavuako, N. Jiang, H. R. Marateb, and D. Farina, "Online mapping of EMG signals into kinematics by autoencoding," *Journal of NeuroEngineering and Rehabilitation*, vol. 15, no. 21, 2018.

[17] K. Bakshi, M. Manjunatha, and C. Kumar, "Estimation of continuous and constraint-free 3 DoF wrist movements from surface electromyogram signal using kernel recursive least square tracker," *Biomedical Signal Processing and Control*, vol. 46, pp. 104–115, 2018.

[18] W. Wang, B. Chen, P. Xia, J. Hu, and Y. Peng, "Sensor fusion for myoelectric control based on deep learning with recurrent convolutional neural networks," *Artificial organs*, vol. 42, no. 9, pp. 272–282, 2018.

[19] T. Bao, A. Zaidi, S. Xie, and Z. Zhang, "Surface-EMG based wrist kinematics estimation using convolutional neural network," in *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 1–4, IEEE, 2019.

[20] T. Bao, S. A. Raza Z., S. Xie, P. Yang, and Z. Zhang, "A CNN-LSTM hybrid framework for wrist kinematics estimation using surface electromyography," *arXiv*, vol. 2, 2020.

[21] A. Phinyomark and E. Scheme, "EMG pattern recognition in the era of big data and deep learning," *Big Data and Cognitive Computing*, vol. 2, no. 3, p. 21, 2018.

[22] C. Wang, W. Guo, H. Zhang, L. Guo, C. Huang, and C. Lin, "sEMG-based continuous estimation of grasp movements by long-short term memory network," *Biomedical Signal Processing and Control*, vol. 59, 2019.