BIOPOINT: SINGLE-SITE, MULTI-SENSOR COMPOUND GESTURE RECOGNITION

Félix Chamberland, Xavier Isabel, Evan Campbell, Gabriel Gagné, Benoit Gosselin, Erik Scheme, Gabriel Gagnon-Turcotte †, Ulysse Côté-Allard †

ABSTRACT

EMG-based gesture recognition tasks have received a lot of attention in recent years, mostly focused on multi-channel EMG sensors, leading to issues in ease of use and computational requirements of such systems. The present study leveraged the BioPoint, a smartwatch-like device, to proceed to multi-sensor deep-learning based gesture recognition. 3 hand gestures in 3 wrist orientations were targeted by measuring the EMG, PPG and IMU waveforms on able-bodied subjects (n=10). Preprocessing and feature extraction allowed the modalities to be used in a two-head neural network trained for the simultaneous classification of hand gestures and wrist rotation. During evaluation, the model obtained an average classification accuracy for hand gestures of $83.5 \pm 12.4\%$ and $94.3 \pm 9.7\%$ for wrist position. Overall, this study showed the potential of single-site, multi-sensor approaches for compound gesture recognition.

INTRODUCTION

Upper extremity musculoskeletal disorders, motor impairments, and amputations profoundly affect millions of individuals worldwide, leading to diverse challenges and significant lifestyle changes [1]. Myoelectric prostheses represent a significant advancement in upper-limb assistive technology to address these issues. These sophisticated robotic devices harness electromyographic (EMG) signals from the muscles in the amputated limb to enable control. Recent progress in robotics has significantly enhanced the capabilities of these prostheses, particularly in achieving various grasp shapes [2]. This enhanced functionality owes much to the integration of multi-channel surface EMG (sEMG) sensor setups, which allow for high-accuracy recognition of multiple concurrent gestures [3], [4]. However, such advancements have primarily been confined to laboratory environments [5].

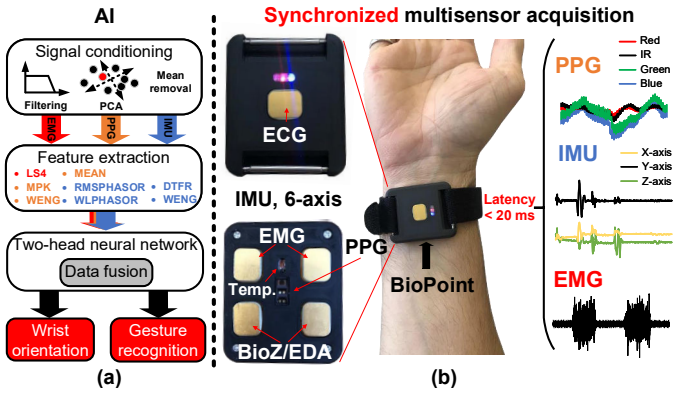

Fig. 1: Concept: (a) The BioPoint provides low-latency synchronized data from up to 6 sensors, (b) the PPG, IMU and EMG signals are processed by a two-head neural network to extract wrist orientation and gesture.

Despite their impressive performance in gesture recognition, multichannel EMG systems face several limitations, including increased complexity, cost and computational requirements, and tend to reduce comfort due to their size [6]. These drawbacks can substantially affect their practicality and user experience.

Moreover, effective prosthetic control encompasses more than just the recognition of diverse grasping actions; it crucially includes the management of wrist dynamics, integral for performing a wide range of activities of daily living [7]. Despite its importance in reducing compensatory movements, which can be a cause for prosthesis rejection [8], the concurrent modulation of grip and wrist motions remains a significant challenge. Studies such as those conducted by Connan et al., exploring the online myocontrol of combined hand and wrist actions, exemplify efforts to overcome these obstacles [9]. Nevertheless, the intricate nature of these challenges underscores the necessity for further innovation and research in signal acquisition, processing, and control algorithms to achieve truly seamless and intuitive control of both grasp and wrist motion simultaneously to enable users to perform a wider array of tasks effortlessly.

Consequently, this work introduces the simultaneous classification of grasp and wrist movements from a single-point, single-channel EMG device as shown in Figure 1(a). This advancement, made possible by the integration of an inertial measurement unit (IMU) and a photoplethysmograph (PPG) within the EMG device, leverages a multi-head deep neural network to perform sensor fusion. The proposed approach is designed to offer a more streamlined, efficient, and user-friendly solution, whilst providing high-level functionality and ultimately improving user experience in prosthetic control.

†U. Côté-Allard and G. Gagnon-Turcotte may be considered co-last authors.

F. Chamberland is the corresponding author (Email: felix.chamberland.1@ulaval.ca). F. Chamberland, X. Isabel, G. Gagné, B. Gosselin, and G. Gagnon-Turcotte are with the Dept. of ECE, Laval University, Québec, Canada. E. Campbell and E. Scheme are with the Department of ECE, University of New Brunswick, Fredericton, NB, Canada, E3B 5A3. U. Côté-Allard is with the Dept. of Technology Systems, University of Oslo, Oslo, Norway.

METHODS

System Hardware

Figure 1(b) illustrates the collection of EMG, PPG, and IMU data conducted throughout the study using the BioPoint [10], a compact and wireless device with the capability of simultaneously recording and streaming EMG, ECG, PPG, IMU, BioZ/EDA, and skin temperature data. Within the context of this study, we focused on three out of the six modalities provided by BioPoint: PPG (Blue, Green, Red, and Infrared) sampled at 50 Hz, a 6-axis IMU sampled at 100 Hz, and EMG sampled at 2 kHz. Figure 1 presents the BioPoint device alongside examples of data captured for various gestures. An important advantage offered by the BioPoint for this study is the synchronization of all sensors using a common clock, eliminating the need for additional processing as all modalities are inherently aligned with each other.

Pre-processing

During the post-acquisition pre-processing phase, the raw data obtained from EMG, PPG, and IMU sensors undergo initial standardization filtering. This filtration process involves removing the mean and adjusting signal magnitudes based on their respective standard deviations, which are derived from the subject's training data. This step ensures a consistent range of values across all three modalities. Additionally, for EMG data, a 60 Hz notch filter is applied to attenuate powerline interference, followed by a 20-450 Hz band-pass filter to eliminate motion artifacts and component noise. PPG data undergoes further processing with a 0.66-3 Hz band-stop filter to mitigate cardiac pulse influences. Moreover, as suggested in [11], a Principal Component Analysis (PCA) is conducted specifically on the Infrared and Red channels. These channels are targeted due to their longer wavelengths, facilitating better penetration into the arm tissue and heightened sensitivity to changes in arm geometry. Regarding IMU data, the approach directly utilizes the standardized signals' 3-axis accelerometer data.

In the context of real-time human-computer interaction, latency is a crucial consideration. To optimize system usability while minimizing input delay, selecting an appropriate window size for data processing is essential. While longer window sizes are associated with enhanced performance in myoelectric control [12], research suggests that the optimal window size to mitigate input latency falls within the range of 150 to 250 ms [13]. Consequently, this study adopts a 200 ms window for classification, supplemented with a 20 ms incremental update, to effectively capture the dynamic nature of the sensor data.

Sensor Fusion Classification Algorithm

A preliminary feature selection was performed for PPG and IMU on pilot subjects to determine appropriate features for each of the modalities. Selection was performed on a per modality basis with the criterion of individual classification accuracy. Subsequently, maximum (MPK), wavelet energy (WENG), and mean (MEAN) features were selected for PPG, and waveform length phasor (WLPHASOR), discrete Fourier transform representation (DTFR), wavelet energy (WENG), and root mean square phasor (RMSPHASOR) features were selected for IMU. The EMG gesture recognition literature was consulted and the LS4 feature set was used, composed of l-score (LS), maximum fractal length (MFL), mean squared ratio (MSR), and Willison's amplitude (WAMP) [14].

The feature vectors are concatenated to form a single input vector, which is then fed into a fully connected neural network. This network comprises three hidden layers, containing 64, 128, and 64 neurons respectively. For each layer, we apply batch normalization, which is then followed by the Scaled Exponential Linear Unit (SELU) activation function as the non-linearity function [15]. Additionally, we employ Alpha Dropout (with a rate of 0.5) after each activation function to reduce overfitting and improve robustness under different input distributions, as suggested by the SELU activation's design principles [15].

The network utilizes a dual fully connected output layer strategy (dual heads), tailored for multitasking. One layer focuses on predicting the three grasps, while the other is dedicated to the wrist movements. The softmax function is applied as a the final layer of non-linearity for both heads.

AdamW [16] is employed with a learning rate of 0.01 to optimize the weights of the network. Learning rate scheduling is used with a step size of 5 epochs and a decay factor (gamma) of 0.1, to adjust the learning rate during training. Furthermore, to prevent overfitting and ensure generalization, we apply early stopping, using 10% of the training dataset as a validation set and a patience parameter set to 30 epochs. Finally, the loss is calculated using cross-entropy, assigning equal importance to both grasp and wrist movement predictions.

Data Collection

In this study, data collection involved ten able-bodied participants, including four women and six men, with ages ranging from 21 to 59 years (Mean: 32.4, SD: 14.8 years). Four participants had no prior experience with biosignal-based collection, and all were new to IMU and PPG data collection methods. The data acquisition protocol received approval from the ethics committee for sectorial research in readaptation and social integration of the *CIUSSS de la Capitale-Nationale* (project 2023-2639). Ethical compliance was ensured through obtaining informed consent from all participants.

Each participant was instructed to execute 3 hand gestures (Neutral grasp (NG), Open hand (OH) and Power grip (PG)) in sequence for three wrist positions: supination (Su), neutral wrist (NW), and pronation (Pr) as shown in Figure 2. Each gesture was recorded for six repetitions lasting 5 seconds each. Only the isometric portions of the contraction were utilized to train the classifier.

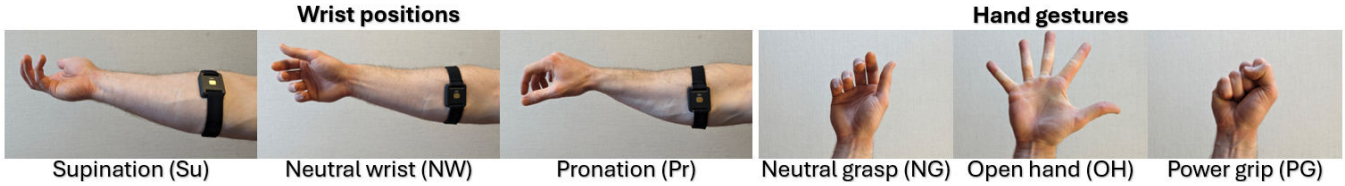

Fig. 2: Gesture set collected during the experiment composed of three hand gestures taken in each of the three wrist positions. The figure illustrates the positioning of the BioPoint sensor on the participant's forearm, located between the brachioradialis and flexor carpi radialis, distal to the elbow joint.

RESULTS AND DISCUSSION

Figure 3 showcases an example of how raw signals from EMG, IMU and PPG differ when comparing different wrist positions and different hand gestures. Neutral grasp, which is used during transitions, has the lowest EMG activity and generally keeps the IMU and PPG signals stable. Open hand causes a slight rise in EMG activity, which is explained by the activation of extensor muscles located on the opposite side of the sensor placement. However, IMU signals have high frequency oscillations similar to the ones obtained with the power grip which have the highest EMG activity due to the recruitement of flexor muscles located right under the sensor. To differentiate between wrist positions, EMG visually seems to play a much smaller role compared to IMU and PPG. The DC levels of the IMU are quite different between the orientations and the shape of the PPG signals are consistent for the same wrist position. These two sensors provide reliable metrics to differentiate amongst the wrist rotations.

Further analysis through feature extraction and neural network training was conducted to obtain quantitative results. The confusion matrix of figure 4(a) shows the results for combined hand grasping and wrist rotation. Both network's heads had to be accurate for the sample to be compiled in the good classification diagonal. One result of note is the generally lower accuracies for the neutral grasp compared to the other active hand gestures. It could be explained by users activating their muscles to maintain pronation or supination which are the two lowest wrist positions for this hand grasp.

Next, by isolating the hand grasps and wrist rotations, classification accuracy with an average of $83.5 \pm 12.4\%$ was obtained for the hand grasps and an average of $94.3 \pm 9.7\%$ for the wrist positions. Figure 4(b) shows the spread of the distribution across the 10 users. Hand grasp was harder to classify due to its reliance on EMG signals which was taken only at a single site far away from the hand. It is to note that the variation in performance across users is significant and the average is skewed downwards by less experienced users. On the other hand, wrist rotations had much higher accuracies except for one outlier user as denoted by the dot under the boxplot at around 67%. Wrist orientations were easier for new users to perform as they do not require the same level of muscular control and practice as finger gestures. Moreover, IMU and PPG sensors were very sensitive to those rotations.

Overall, this experiment showed that separating the hand and wrist motions using multiple sensors on a single-site already obtains great results. Therefore, performing a similar compound movement experiment using sensor fusion and multi-site EMG is a promising next step to hopefully enhance control of the newer generation of fully articulated prosthesis.

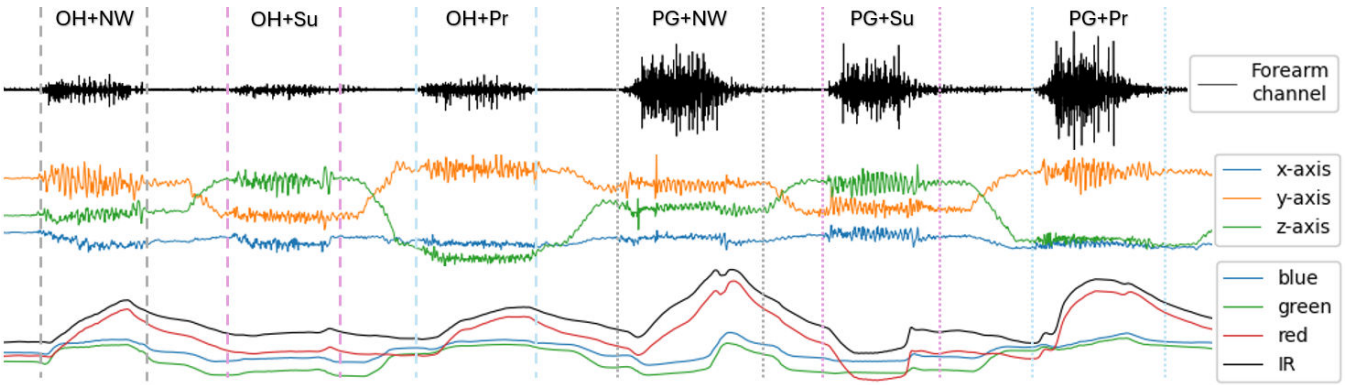

Fig. 3: A sequence was performed to showcase how the EMG, IMU and PPG raw signals affected by hand and wrist motion. 6 combinations of active hand gestures and wrist positions were performed. Transition periods were used to change wrist orientation while keeping the hand in a neutral grasp (passive).

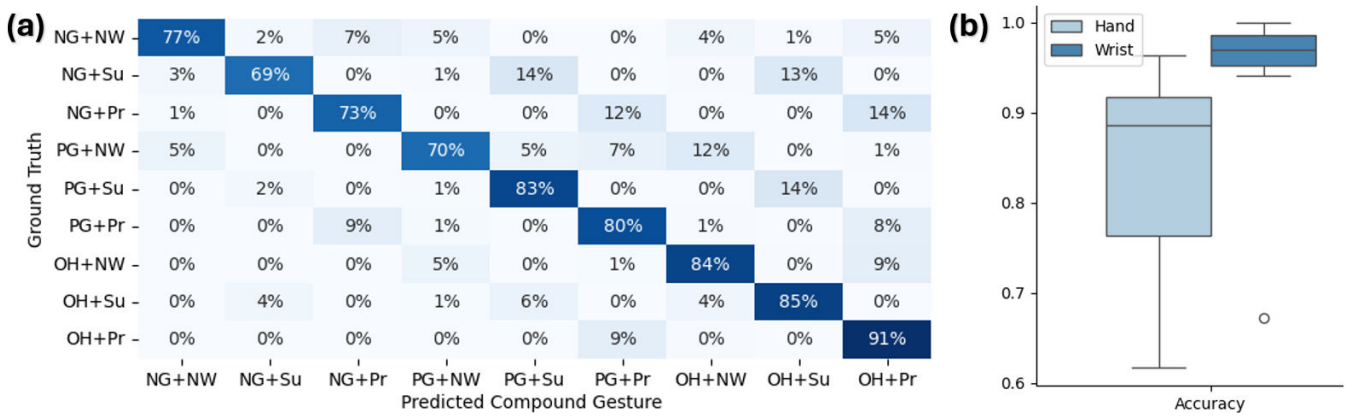

Fig. 4: (a) shows the confusion matrix for the grasps and wrist motions considered simultaneously as predicted by the dual-head deep neural network utilizing sensor fusion in the proposed classification pipeline. (b) provides the box plots, showcasing the distribution of the participants' accuracies for both grasp (hand) and wrist motions predicted separately.

CONCLUSION

In this study, the BioPoint, a single-site multi-modal acquisition device, was used to implement a two-head neural network for simultaneous classification of hand gestures and wrist rotations, laying the groundwork for integrating wrist dynamics into gesture recognition systems. Following preprocessing of EMG, IMU, and PPG waveforms, features supported by literature and experiments were extracted. The network's performance was evaluated with common offline metrics and showed near-perfect performance for wrist rotations and high accuracy for hand gestures. While these results are promising, further experiments should capitalize on online classification and on evaluating the system's effectiveness with upper-limb amputees, both of which are crucial aspects for the practicality of myoelectric prostheses control.

REFERENCES

[1] K. Ziegler-Graham, E. J. MacKenzie, P. L. Ephraim, T. G. Travison, and R. Brookmeyer, "Estimating the prevalence of limb loss in the united states: 2005 to 2050," *Archives of physical medicine and rehabilitation*, vol. 89, no. 3, pp. 422–429, 2008.

[2] J. T. Belter, J. L. Segil, and B. SM, "Mechanical design and performance specifications of anthropomorphic prosthetic hands: a review," *Journal of rehabilitation research and development*, vol. 50, no. 5, p. 599, 2013.

[3] F. Chamberland, É. Buteau, S. Tam, E. Campbell, A. Mortazavi, E. Scheme, P. Fortier, M. Boukadoum, A. Campeau-Lecours, and B. Gosselin, "Novel wearable hd-emg sensor with shift-robust gesture recognition using deep learning," *IEEE Transactions on Biomedical Circuits and Systems*, 2023.

[4] U. Côté-Allard, G. Gagnon-Turcotte, F. Laviolette, and B. Gosselin, "A low-cost, wireless, 3-d-printed custom armband for semg hand gesture recognition," *Sensors*, vol. 19, no. 12, p. 2811, 2019.

[5] E. Campbell, A. Phinyomark, and E. Scheme, "Deep cross-user models reduce the training burden in myoelectric control," *Frontiers in Neuroscience*, vol. 15, p. 657958, 2021.

[6] M. Tavakoli, C. Benussi, and J. L. Lourenco, "Single channel surface emg control of advanced prosthetic hands: A simple, low cost and efficient approach," *Expert Systems with Applications*, vol. 79, pp. 322–332, 2017.

[7] 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.

[8] A. J. Metzger, A. W. Dromerick, R. J. Holley, and P. S. Lum, "Characterization of compensatory trunk movements during prosthetic upper limb reaching tasks," *Archives of physical medicine and rehabilitation*, vol. 93, no. 11, pp. 2029–2034, 2012.

[9] M. Connan, R. Kõiva, and C. Castellini, "Online natural myocontrol of combined hand and wrist actions using tactile myography and the biomechanics of grasping," *Frontiers in Neurorobotics*, vol. 14, p. 11, 2020.

[10] G. Gagnon-Turcotte, U. Côté-Allard, Q. Mascret, J. Tørresen, and B. Gosselin, "Photoplethysmography-based derivation of physiological information using the biopoint," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2023, pp. 1–5.

[11] Y. Ruan, X. Chen, X. Zhang, and X. Chen, "Principal component analysis of photoplethysmography signals for improved gesture recognition," *Frontiers in Neuroscience*, vol. 16, p. 1047070, 2022.

[12] 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.

[13] L. H. Smith, L. J. Hargrove, B. A. Lock, and T. A. Kuiken, "Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 19, no. 2, pp. 186–192, 2010.

[14] A. Phinyomark, R. N. Khushaba, and E. Scheme, "Feature extraction and selection for myoelectric control based on wearable emg sensors," *Sensors*, vol. 18, no. 5, p. 1615, 2018.

[15] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-normalizing neural networks," *Advances in neural information processing systems*, vol. 30, 2017.

[16] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.