



Understandable time frame-based biosignal processing

Hamed Rafiei , Mohammad-R. Akbarzadeh-T ^{*}

Department of Electrical Engineering, Center of Excellence on Soft Computing and Intelligent Information Processing, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran

ARTICLE INFO

Keywords:

Artificial intelligence
Biosignal processing
Classification
Explanation
Time series

ABSTRACT

The explainability of biological time series poses considerable challenges regarding signal multiplicity, high volatility, nonstationarity, and noisiness in pursuit of understanding human intentions and conditions. These challenges often arise since data points in the time series are inherently unexplainable and need complex models for proper processing. Here, we propose data frames as a primary information unit. Specifically, the proposed biosignal time frame (BioTF) series incorporates data frames inspired by candlestick components from financial data analysis, such as starting, highest, lowest, and ending values (SHLE). We implement BioTF on four benchmarked biosignal classification tasks, including electromyography (EMG), high-density surface electromyography (HD-sEMG), electroencephalography (EEG), and electrocardiogram (ECG). We study various time frame lengths, components, network activation functions, and architectures for these instances. The bio time-frame representation shows similar patterns, technical analysis, and results to financial data analysis, offering an exciting analogy between these two domains. Compared with several point-based strategies, the proposed BioTF improves temporospatial explainability and achieves as much as 7% improved classification due to the reduced complexity by extracting intuitive features in the proposed frame-based representation. The proposed BioTF furthermore leads to competitive results using simpler networks with as much as four times faster end-to-end training and lighter frame-based feature extraction after-step training. The proposed method enables lightweight and transparent implementations of AI recommendation systems for expert manipulation and trustworthy medical translations of bio time series. The proposed SHLE representation is general and could be extended towards more detailed signal representations.

1. Introduction

Biosignals represent human states. Hence, they have extensive utility in devices closely interacting with humans, such as health monitoring and bionic limb control systems. Deep learning architectures have provided high classification and regression accuracy for such applications in recent years. However, the interpretability and reliability of biological signals for condition/intention detections and their associated decisions continue to daunt the real-world applicability of such architectures [1,2].

Explainable artificial intelligence (XAI) aims to explain the logic behind data-driven models. Explanation methods can be categorized according to the numerical/categorical, pictorial, textual, and time series data they process. They can also be classified according to the intelligent architecture they employ. For instance, several explanation methods use deep learning architectures, such as gradient-based, relevance score-based, attention-based, generation-based, and node-based

[3]. Specifically, layer-wise propagation computes a relevance score by multiplying each layer's input and gradients [4]. In [5], a class activation mapping (CAM) shows the image regions that convolutional neural networks (CNNs) use to identify a specific class. These methods often explain decisions for image and textual information. However, time series explanations have much to achieve since, unlike pictures and text, humans cannot intuitively visualize and perceive them at a glance. These explanations become even more challenging in long-time durations [6], especially in end-to-end deep architectures.

Few works have recently addressed time series explanations. In [7], a CAM strategy similar to [4] visualizes the series region leading to labels. In [8], CNNs are visualized at different layers' generated abstractions for time series classification and regression. In [9], counterfactual observations are replaced by actual ones to determine observation importance by measuring the expected change of predictions. A few other works consider region visualization instead of visualizing data points for better interpretation. In [10], the network dissection method determines the

* Corresponding author.

E-mail addresses: rafei.hamed@mail.um.ac.ir (H. Rafiei), akbazar@um.ac.ir (M.-R. Akbarzadeh-T).

alignment between the CNN's unit response and a dense segmented image data set. In [11], CNN is compared in identifying the critical regions identified by clinicians. Theissler et al. in [12] have provided an extensive review of the recent works on XAI for time series classification.

Explainability is a central concern in medical applications because clinicians should understand the reason behind model decisions to ensure they are determined based on essential and relevant features. Current works often visualize causalities on a limited number of resampled data points due to input representation constraints. In [13], explanations show that the model uses the physiological information of EMG to determine finger joint angles. Recently, Cho et al. [3] provided a regional explanation for EEG signals by clustering sequences that highly activate the internal nodes. The resulting representative sequences then offer insight into the overall shape of the discriminative patterns. In [14], the introduced LAXCAT method determines the critical time intervals and spatial regions of multivariate time series classification. In [15], spectral, spatial, and temporal explanations of classified EEG time series are provided for seizure detection. A gradient-weighted class activation mapping (Grad-CAM) in [16] validates a CNN's ability to learn biologically plausible features of expert game player brains. The Shapley additive explanation method in [17] reveals the reasons behind a hybrid CNN-LSTM network's decision regarding ECG arrhythmia. The explanation in [18] shows how a network detects COVID-19 using the ECG ST segment. In [19], the Grad-CAM visualization method explains parts of the ECG signal that a self-organized-operational neural network heeds. While these methods provide valuable insights, a crucial remaining step is to fill the gap between technical explanations and true understandability for diverse stakeholders, including non-technical clinicians and patients [20]. In other words, we must focus on developing methods that reveal the underlying mechanisms of AI decisions and present this information in a clear, accessible manner that facilitates genuine comprehension and informed decision-making across various levels of technical expertise.

Besides end-to-end data processing, some works perform pre-processing, feature extraction, dimension reduction, and decision-making separately. Noise and motion artifacts generally corrupt the biological signal acquisition process. So, pre-processing, including filtering and data scaling, is often performed after biosignal acquisition to remove the non-informative data and enable the accurate identification of the signal's meaning [21]. After pre-processing, features provide informative data for identifying conditions and intentions. These features are presented in time, frequency, or time-frequency domains [22–24]. Time-domain features estimate time-domain parameters of biosignals, while frequency-domain features extract spectral information. The time-frequency representation takes advantage of both domains to reveal the signal frequency content over time. These features compete in computational cost, complexity, and information contact. They require more computations and hyperparameter tuning to extract more information. For example, the zero crossing rate (ZC) feature from the time domain determines how many times the signals cross the zero axis with a conditional term, while the time-frequency wavelet transform feature captures changes in amplitude and phase in different frequencies over time [21]. Since there may be many features, including some with irrelevant information, dimension reduction methods are often applied to 1) keep the most informative and 2) increase data visualization ability [25]. These features are then used for decision-making. In [26], the DeepForest strategy explains finger force modeling based on high-density surface electromyography features. The

more difficult question is which of these most helpful biosignal features are also more straightforward to extract, more intuitive to understand, and more informative to classify.

The above dilemma could be addressed by changing how we represent data at the first stage. The present work underscores the fact that a suitable data representation could be instrumental in revealing the main patterns within the signal. As such, the causes of classification or regression decisions could also be more understandable. Such causalities are essential for sustainable medical and health developments [13]. Yet, to our best knowledge, no previous works have directly addressed biosignals' temporal representation for interpretation purposes.

This research proposes a new representation of biological time series to make explanations more understandable. We then propose to process the time series directly in its temporal data frame domain. The approach represents each channel of window size n of biosignals by its high, low, starting, and ending values in time frames of length l , creating four series of length n/l . This representation is similar to the candlestick representation common in finance, expecting to achieve a comparable level of interpretability for biological time series. Compared to traditional bio-signal feature extraction, partitions here have added values, where features are extracted in three-dimensional (amplitude, time, and frame) rather than two-dimensional (amplitude and time) space. This representation is then processed by neural networks for classification purposes in our first set of experiments on three bio datatypes. In other words, the networks are driven to classify based on the time frames instead of the data points. Furthermore, explanations of the proposed frame structure facilitate understanding the intuitive patterns that affect the networks' decisions. In our second set of experiments, linear discriminant analysis (LDA) classifies HD-sEMG into 34 hand gestures with time frame- and data point-based features.

In summary, the main contributions of this paper are in two folds, as follows:

- Time frame representation of biosignals: By utilizing the common visualization approaches more effectively, we readily illustrate the latent patterns of a biosignal. This representation contrasts the usual partitioning and segmentation methods that break series into data point subseries [27,28]. Instead, building upon the explained similarity between financial and biosignal processing, we represent each partition into a time-amplitude-embedded feature vector. Hence, similar to financial knowledge, clinical knowledge could be directly used to extract meaningful hand-crafted patterns.
- Time frame processing and explanation: The time frame concept leads to a simple, highly explainable decision system, mainly because the networks, from the beginning, learn to classify the represented signals that are more understandable to humans. In fact, we provide an understandable explanation, in contrast to previous endeavors that make the explanation more understandable by clustering the highly activating internal nodes [3] or focusing on the attention matrix [13].

Extensive experiments on EMG, EEG, and ECG signals are performed under various time frames to show the proposed BioTF's superiority in terms of accuracy, explanation, and training computational cost. The classification explanations are provided by node activation and local interpretable model-agnostic explanations (LIME) techniques. Furthermore, the HD-sEMG signal classification is compared with state-of-the-art results.

The remainder of the paper is organized as follows. Section II introduces the proposed BioTFs. Simulation results, including dataset, experiments on networks' activation functions, layers, input components, changing frame slide, explanation, and computational time, are provided in Section III. Section IV discusses the results, and Section V concludes the paper. Finally, technical analysis, dataset and architecture details, feature sets, and explanation parameters.

2. The proposed biosignal processing based on data time frames (BioTF) method

The proposed method is described in this section by biological and financial time series analogies explanation, time frame conversion, and processing subsections.

2.1. The analogy explanations between biological and financial time series

Besides feature extraction, signal representation provides a proper basis for further analysis, similar to what happens in finance. Prices are the amounts payable or compensated by one party to another in return for goods or services. Production cost, supply, and demand mainly affect prices, sometimes making them unpredictable and volatile. Japanese rice traders first invented candlesticks showing the market's open, high, low, and close prices throughout a specific period to observe the market trend more easily. This kind of representation also removes the necessity for an in-depth understanding of market microstructure, making it easier to recognize simple, handcrafted patterns. Since then, various market strategies have been used to predict and trade [29].

The following examples illustrate the significance of time frame representation and related candlestick patterns in finance and biosignals.

Example 1. Time frame representation visually provides a better sense than ask-bid tick data. For better comparison, see Fig. 1, which shows a sine function with added noise. Detecting the decline data point from the original signal (top signal) is not trivial. However, when we use a time frame with 20 samples per candlestick, we observe a decreasing pattern when the open value of a candle is below the previous candle's closing value (bottom signal). On the other hand, biosignals illustrate human conditions. These biosignals are often recorded at high frequency and include higher frequency noises that humans can not comprehend at a simple glance. However, time frame representation could be helpful. Fig. A.1 illustrates the representation of biosignals using time frames.

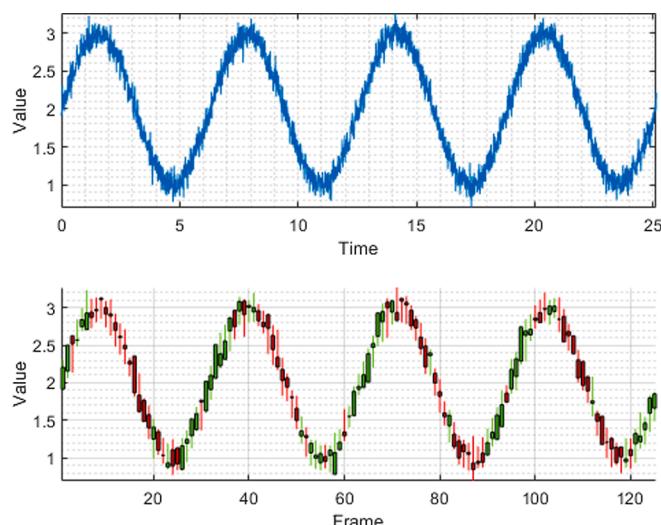


Fig. 1. A simple time series illustration into 20 sample-length time frames.

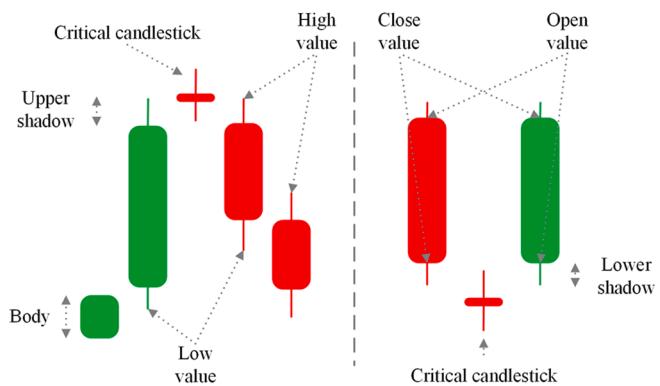


Fig. 2. Bearish Doji Star (Left) and Bullish Abandoned Baby (Right), where the green and red candles indicate increasing and decreasing signals, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Example 2. The Doji Star Bearish pattern shows downward reversal data points, identified by a Doji candle (Fig. 2). In fact, there is a Bearish pattern at the end of the long upward prices and the beginning of the downward prices. The pattern is said to have high reliability when the Doji candle appears above the close price of the second candle, with the fourth candle opening below the Doji star. Traders often wait for the closing of the second candle to open a short position (short class). On the other hand, in biosignals, such time frame patterns could identify a type of arrhythmia. For example, the pattern and its time frame characteristics (body and shadows) could be used to recognize heart disease based on ECG signal peaks [30].

Example 3. The Abandoned Baby Bullish pattern represents an upward reversal movement and includes three consequent candles (Fig. 2). This pattern happens when a Doji candle is abandoned by two adjacent long downward and upward candles. At this time, a potent long position arises (long class). In the sense of biosignals, these patterns could be applied to determine a trigger. For example, bionic hands often work based on EMG signals and need a trigger to start or stop a movement. In this case, such patterns reveal the stop and reversal movements.

In short, we believe that bio and financial time series have two main analogies, including 1) both time series are nonstationary. This property makes them almost unpredictable because their characteristics constantly change over time. For example, an increase in a stock price could be followed by a sudden dropping. Also, speed, strength, and range of voluntary motions affect the amplitude and spectral frequency of EMG signals and change due to muscle fatigue [31]. And 2) we should be aware that we have no control over these series. We can observe the financial series and decide to buy, sell, or hold, but we cannot set. In the same way, we observe humans' biosignals and recognize their condition or intention for further notification of a state or to control bionic limbs. We cannot order to generate a unique biosignal.

We further investigate if there are class-specific patterns in biosignals. To this aim, we apply 90 famous financial patterns to a biosignal dataset. EMG signals are very noisy and often recorded at high frequencies to control bionic limbs, making them suitable for this investigation.

Rickshaw Man's financial pattern shows that a specific change in signal is not expected. This pattern's starting and ending values are very close with long shadows. Investigating this pattern within the EMG signal shows that this pattern could be used to highlight the Rest gesture of the bionic hand (Fig. B.1). So, one could conclude that there is a potential to find specific patterns for each class.

Also, several technical analyses provide insight into the market. From the perspective of technical analysis of financial data, several works in EMG-controlled bionic hand control perform a similar manner

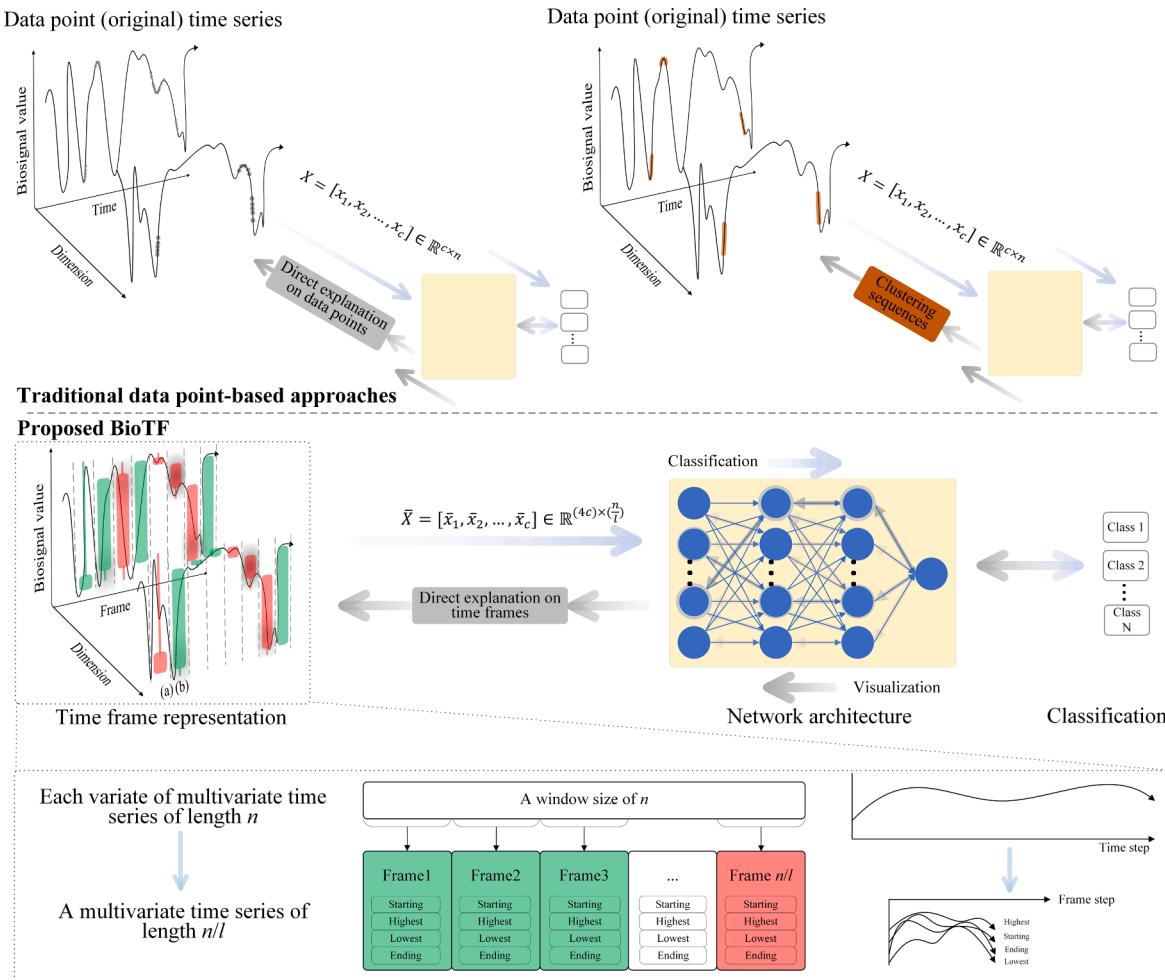


Fig. 3. Traditional data point-based approaches of processing time series (top), where solid black lines show the original biosignals and explanations highlight the important data points (gray spots) or sequences (orange lines). In the proposed BioTF (bottom), the green and red candlesticks represent the characteristics of their corresponding time frames. For example, the red candle (a) has a long upper shadow and ends at a lower value than its starting, while the green candle (b) monotonically increases until its ending value without shadows. The gray rectangular highlights show the relative importance of each time frame. Both approaches are assumed to use a similar network (with a difference in input dimension), with the blue and gray lines representing the forward classification and backward visualization passes, darker lines on the weights showing a more effective related pass, and the more highlighted nodes showing higher activation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

for grasp force estimation [32] and the contraction onset determination using the mean absolute value (MAV) of the EMG transient signal [33,34]. Our study also shows that *atr* indicators have a good potential for discriminating the EMG classes, having different levels for each class (Fig. B.2). More analyses are provided in Appendix B.

This brief subsection shows the potential of using time frame biosignals to establish a decision system by machines and experts. Based on the above results, we develop deep learning structures in this work to explore discriminative patterns and analyze them further. More results on applying technical analysis are provided in Appendix C.

2.2. Time frame representation

In comparison to the literature, the proposed method is shown in Fig. 3. We partition temporal data points from each signal channel into

non-overlapping time frames with a specific length. The SHLE values are then determined for the related window (adding signal power dimension is investigated in Appendices). In other words, if we have a multichannel signal chunk $X = [x_1, x_2, \dots, x_c] \in \mathbb{R}^{c \times n}$ with c channels and n data points length of a window, the converted signal is $\bar{X} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_c] \in \mathbb{R}^{(4c) \times \frac{n}{l}}$ where l is the frame length (without loss of generality, in a data stream, the time frame representation progresses when the number of observations in each frame reaches l). For example, the i^{th} SHLE values are calculated for the channel x_k as follows:

$$\bar{x}_k(i) = \begin{bmatrix} x_k((i-l)l+1) \\ \max(x_k((i-l)l+1 : il)) \\ \min(x_k((i-l)l+1 : il)) \\ x_k(il) \end{bmatrix} \in \mathbb{R}^{4 \times 1},$$

$$\text{for } i \in [1, 2, \dots, \frac{n}{l}], k \in [1, 2, \dots, c], \text{ and } \text{mod}(n, l) = 0. \quad (1)$$

2.3. Time frame-processing

The network architecture uses the time frame series to learn related classes (Fig. 3). The process of classification by the proposed BioTF could be formulated as follows:

$$Y = f(\bar{X}; \theta), \quad (2)$$

where the mapping function $f(\cdot; \theta) : \mathbb{R}^{(4c) \times (\frac{n}{l})} \rightarrow \mathbb{R}^q$ could be any supervised learning methods with parameters θ such as neural networks and $Y \in \mathbb{R}^q$ is the associated label for the input \bar{X} .

After that, we use traditional visualization techniques to understand how networks use time frames to specify classes. Therefore, explanation $\xi(\bar{X})$ using traditional methods such as layer activation and LIME [35] reveals the importance of candlestick components through time. This measure is then averaged to show candlestick importance, $\bar{x}_k(i) = 1/4 \sum \bar{x}_k(i)$. In other words, the effect of time frames could easily be understood instead of data points. Algorithm 1 presents the pseudocode

of the proposed BioTF for predicting classes over chunk X .

Algorithm 1. Procedure of BioTF.

Input: Biosignal time series chunk X
Output: decision and explanation
Initialization: trained classifier $f(\cdot; \theta)$ and frame length l .
Get the next chunk

Time frame representation of the chunk X using Eq. (1).

$$\bar{X} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_c] \in \mathbb{R}^{(4c) \times (\frac{n}{l})}$$

for $i \in [1, 2, \dots, \frac{n}{l}]$ and $\text{mod}(n, l) = 0$

for $k \in [1, 2, \dots, c]$

$$\bar{x}_k(i) = \begin{bmatrix} x_k((i-l)l+1) \\ \max(x_k((i-l)l+1 : il)) \\ \min(x_k((i-l)l+1 : il)) \\ x_k(il) \end{bmatrix} \in \mathbb{R}^{4 \times 1}$$

Time frame processing $Y = f(\bar{X}; \theta)$, using Eq. (2).

$$f(\cdot; \theta) : \mathbb{R}^{(4c) \times (\frac{n}{l})} \rightarrow \mathbb{R}^q$$

Perform post hoc explanation $\xi(\bar{X})$.

The average on dimensions to show candlestick importance $\bar{x}_k(i) = 1/4 \sum \bar{x}_k(i)$.

3. Results

In this section, we provide several simulations to examine the

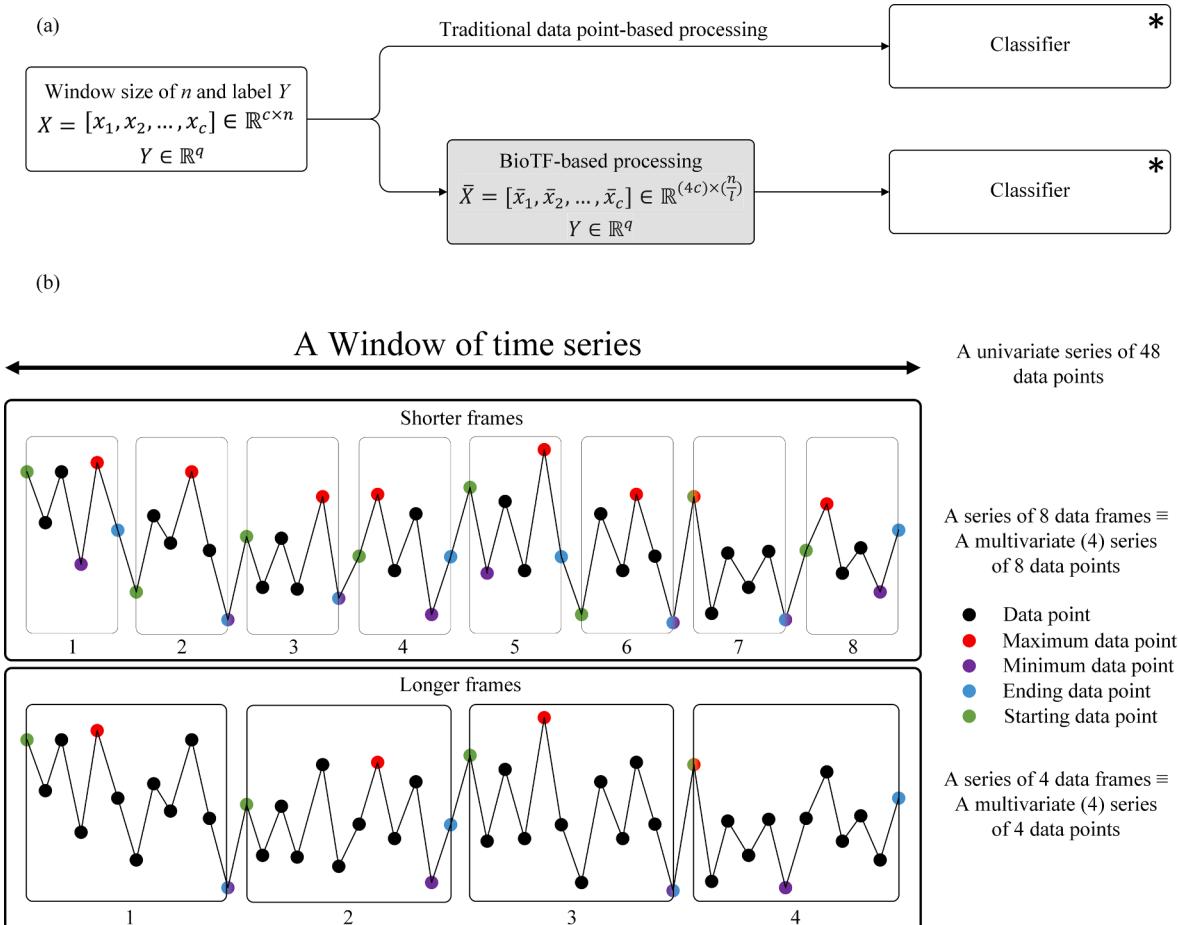


Fig. 4. (a) Settings and conditions of comparisons. * means the same classifier architectures and training. (b) An illustration of BioTF representation with different frame lengths (shorter and longer frames).

proposed BioTF. Each experiment is performed three times with random seeds, and the average values are reported. All the comparisons have the same settings and conditions, as shown in Fig. 4a. We use a series of data points for traditional data point-based classification and a transformation (frames) of the series of data points for BioTF-based classification. For example, we use a window that includes 48 data points, known as original signals, to determine associated labels in traditional classification based on data points (Fig. 4b). We also use the same original signal for the BioTF as our database. But we transform it into a series of frames with widths of 6 and 12 (each data window consists of several consecutive frames by the BioTF). For instance, using a frame width of 6, we have a series of eight ($=48/6$) frames for BioTF-based processing. Overall, traditional signal processing and BioTF use the same window size as a primary step of the classification procedure. All the experiments are conducted on a desktop computer with a 3.4 GHz Quad-Core Intel Core i7 processor, 16 GB of RAM, and an NVIDIA GeForce GTX 1050 graphic card. MATLAB (2022b) software executes the experiments.

3.1. Datasets

The proposed BioTF is applied to several bio datasets, including EMG, EEG, ECG, and HD-sEMG signals. EMG signal shows the muscle's electrical activity in response to a muscle nerve's stimulation. The EMG is often used in bionic wearable robots to determine human intention and motion. Furthermore, EEG signals record brain activities and have several medical uses, such as epilepsy detection. Moreover, ECG presents the electrical activity of the patient's heartbeats. Physicians observe ECGs to detect normal or irregular conditions in patients. More details on the studying datasets are provided in Appendix D.

The summarization of the statistics, given in Table I, shows that the datasets have different characteristics. Results show that the signals are disturbed more from down to the top, causing negative SNR. Furthermore, EMG and EEG have more unpredictable fluctuations because of higher AE. Also, LE results [36] state that signals exhibit chaotic behaviors more from down to top. These datasets are multiple time series (#channel > 1) except ECG.

3.2. Experiment settings

We use LDA (the same as [37]) for the HD-sEMG and neural networks for other datasets. The network architectures are more flexible for containing various tools and connections, making them appropriate for learning procedures in various applications. Although there are complex architectures for specific applications such as EMG [38], EEG [39,40], and ECG [41,42] classifications, we use general architectures to focus on the time frame-based representation advantages. Because the EMG dataset is long and its sampling rate is high, we perform experiments with up to a frame length of $l = 50$, but smaller lengths are used for other datasets. The network architectures and the related hyperparameters are provided in Appendix G.

Table 1
Characteristics of bio datasets.

| Dataset | Channel | RMS | Variance | Skewness | Kurtosis | SNR | AE | LE(min/max) | F _{ins} (min/max) | T = 1 / F _{ins} |
|---------|---------|--------|-----------|----------|----------|-------|-----|---------------|----------------------------|--------------------------|
| EMG | 8 | 0.03 | 0.002 | 0.007 | 3.06 | -13.5 | 1.3 | 200(-1e3/2e3) | 445 /1.05e + 3 | 0.002 /9.5e-4 |
| EEG | 21 | 2.14 | 6.131 | -0.120 | 4.26 | -3.2 | 0.5 | 64(-121/166) | 0.83 /118.77 | 1.2 /0.008 |
| ECG | 1 | 198.11 | 4.96e + 4 | 1.021 | 16.52 | -6.30 | 0.2 | 130(67/171) | 1.25 /49.65 | 0.8 /0.02 |
| HD-sEMG | 256 | 0.1127 | 0.0289 | 0.0754 | 5.1827 | -5.51 | 0.9 | 714(-2e3/1e3) | 6.8 /724.7 | 0.14 /0.001 |

All the metrics are calculated using build-in functions (*rms*, *var*, *skewness*, *kurtosis*, *snr*, *approximateEntropy*, *lyapunovExponent* and *instfreq*) of MATLAB (2022b) software. All the values are averaged over recordings' channels.

3.3. Accuracy results

We compare the time frame-based signal processing results with data points for each dataset. Table 2 shows the results summary of the simulations. No significant superiority in data point-based processing over the time frame exists. As the data is more complex and noisier, results with longer time frames are more competitive to the data point-based processing. In ECG signals, the accuracy significantly decreases due to hiding important points in the frames. However, the accuracy of the ECG signal reaches up to 92.98 % when we use overlapped frames (please see Section III.D.e). Furthermore, the classification accuracy of EMG and EEG signals could reach 89.61 % and 86.8 % using overlapped frames and the Sine activation function, respectively (please see Section III.D.e and III.D.a, respectively). The comparison of maintenance tasks on HD-sEMG signals is provided for data point-based feature sets [37] and frame-based feature sets. Results indicate that a single time frame-based feature with less computation competes with data point-based features. More details on the feature sets are provided in Appendix H.

3.4. Ablation study on accuracy

To support the above results, we continue this line of investigation with more experiments under various conditions. Specifically, we perform experiments on different activation functions, applying normalization, using the dropout layer, and reducing the input dimension. The subsections below provide more details.

3.4.1. Different activation

In these experiments, we change networks' activation functions to the Sine and Sinc functions. We use these activations to examine network classification ability based on the periodic activation functions.

- Sine activation function

In these experiments, we add after the LSTM layer in all networks except the EEG network and replace the "eluLayer" with a periodic activation function as follows:

$$F(z) = \text{Sin}(z), \quad (3)$$

where, z is the input to the layer. Table 3 represents the classification accuracies. Results show that the accuracy on all signals increases using the time frame $l = 5$ except for ECG. Also, the classification accuracy on the EEG dataset reaches the highest value of 86.8 percent in all experiments.

- Sinc activation function

In these experiments, we add after the LSTM layer in all networks except the EEG network and replace the "eluLayer" with a periodic activation function as follows:

$$F(z) = \text{Sinc}(z), \quad (4)$$

Table 2

The comparison of the traditional and time frame signal processing. The best results in this experiment are in bold and all experiments are in italics.

| Dataset | Train (Test) accuracy /Test F1 score over frame length l | | | | | |
|---|--|------------------------|--------------------------------|-----------------|-----------------|-----------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 82.1(81.2)/ 0.81 | 81.8(80)/0.792 | 80.9(79.4)/0.782 | 81(79.8)/0.79 | 80.6(79.2)/0.78 | 79.9(79)/0.78 |
| EEG | 82.1(83)/ 0.825 | 83.1(82)/ 0.835 | 84.5(84)/ 0.83 | 83.6(83)/0.83 | — | — |
| ECG | 80.5(82)/ 0.715 | 82(70)/0.675 | 63(55)/0.49 | 77(72)/0.485 | — | — |
| Test accuracy of Maintenance task on original signal over various feature sets (average on cross-validations and sessions). | | | | | | |
| | D1 [#] | D2 [#] | D3 [#] | D4 [#] | D5 [#] | D6 [#] |
| HD-sEMG [†] | 93.8 | 91.1 | 92.4 | 92.3 | 93.7 | 87.3 |
| | D7 [#] | D8 [#] | F1 ^{\$} | | | |
| | 90 | 90.2 | 93.5 | | | |

* Means no frame has been applied, and the original signal has been processed without BioTF representing.

Data point-based feature set.

\$ Time frame-based feature set.

† Only original signals of maintenance tasks are included from the HD-sEMG dataset.

Table 3

The comparison accuracy of the traditional and time frame signal processing using the Sine activation function. The best results in this experiment are in bold and all experiments are in italics.

| Data | Train (Test) Accuracy /Test F1 score over frame length l | | | | | |
|------|--|--------------------------------|-------------------------|-------------------------|-------------------------|-----------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 70.1 (68.1)/ 0.80 | 82.4 (80.9)/ 0.77 | 81.9 (80.1)/ 0.79 | 80.7 (79.8)/ 0.78 | 79.5 (79.4)/ 0.78 | 80.1 (79)/ 0.77 |
| EEG | 86.5 (85.1)/ 0.84 | 87.8 (86.8)/ 0.86 | 84 (84.6)/ 0.84 | 82.1 (82.9)/ 0.82 | — | — |
| ECG | 84.5 (83.0)/ 0.83 | 63.3 (56.0)/ 0.56 | 63.6 (50.5)/ 0.46 | 78.5 (62.6)/ 0.56 | — | — |

* Means no frame has been applied, and the original signal has been processed without BioTF representing.

Table 4

The comparison accuracy of the traditional and time frame signal processing using the Sinc activation function. The best results in this experiment are in bold and all experiments are in italics.

| Data | Train (Test) Accuracy /Test F1 score over frame length l | | | | | |
|------|--|--------------------------------|--------------------------------|-------------------------|-----------------------|-------------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 53.7 (53.4)/ 0.49 | 82.4 (82.5)/ 0.81 | 82.6 (82.2)/ 0.81 | 82.1 (81.4)/ 0.79 | 80.7 (80)/ 0.78 | 65.1 (64.1)/ 0.60 |
| EEG | 43 (42.9)/ 0.35 | 52.1 (51.4)/ 0.51 | 57.6 (57.0)/ 0.55 | 52.1 (51.9)/ 0.45 | — | — |
| ECG | 71.9 (69.2)/ 0.69 | 78.7 (48.5)/ 0.33 | 49.2 (50.6)/ 0.36 | 50.3 (49.6)/ 0.36 | — | — |

* Means no frame has been applied, and the original signal has been processed without BioTF representing.

where z is the input to the layer. Table 4 shows the related results. Results show that the accuracy of all signals increases using the time frame approach except ECG.

3.4.2. Adding layers

After examining periodic activations, we evaluate the network performance when adding normalization and dropout layers.

- *Adding normalization layer*

A normalization layer is added after the LSTM layer in EMG and ECG networks to improve the network's training convergence and reduce the hyperparameter selection sensitivity. The EEG network already has the

Table 5

The comparison accuracy of the traditional and time frame signal processing by adding a normalization layer. The best results in this experiment are in bold and all experiments are in italics.

| Data | Train (Test) Accuracy /Test F1 score over frame length l | | | | | |
|------|--|--------------------------------|-------------------------|--------------------------------|-------------------------|-------------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 86.9 (86.3)/ 0.86 | 88.6 (88.3)/ 0.87 | 86.1 (86.2)/ 0.85 | 86.1 (86.1)/ 0.85 | 85.7 (85.5)/ 0.85 | 83.9 (83.1)/ 0.83 |
| ECG | 50(50)/ 0.33 | 81.4 (80.9)/ 0.80 | 76.8 (72.6)/ 0.72 | 92.6 (85.9)/ 0.86 | — | — |

* Means no frame has been applied, and the original signal has been processed without BioTF representing.

Table 6

The comparison accuracy of the traditional and time frame signal processing by adding a dropout layer. The best results in this experiment are in bold and all experiments are in italics.

| Data | Train (Test) Accuracy /Test F1 score over frame length l | | | | | |
|------|--|----------------------------------|-------------------------|-------------------------|-----------------------|-------------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 80.4 (80.3)/ 0.79 | 81.1 (80.5)/ 0.77 | 79.9 (79.7)/ 0.78 | 80 (79.4)/ 0.78 | 80 (79.3)/ 0.77 | 78.8 (78.8)/ 0.77 |
| ECG | 69.56 (69.9)/ 0.33 | 75.83 (74.98)/ 0.75 | 57.2 (49.8)/ 0.43 | 68.2 (58.8)/ 0.57 | — | — |

* Means no frame has been applied, and the original signal has been processed without BioTF representing.

Table 7

The comparison accuracy of the traditional and time frame signal processing by adding signal energy dimension. The best results in this experiment are in bold and all experiments are in italics.

| Data | Train (Test) Accuracy /Test F1 score over frame length l | | | | | |
|------|--|--------------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 81.5 (81.2)/ 0.81 | 89.1 (88.2)/ 0.88 | 87.5 (87.0)/ 0.86 | 86.9 (86.7)/ 0.86 | 86.8 (86.1)/ 0.86 | 84.8 (84.3)/ 0.84 |
| EEG | 83.6 (83)/ 0.825 | 58 (57.4)/ 0.39 | 68.1 (67.2)/ 0.59 | 74.1 (74.0)/ 0.71 | — | — |
| ECG | 80.5 (82)/ 0.715 | 79.3 (75.8)/ 0.75 | 78.5 (70.4)/ 0.69 | 92.7 (75.4)/ 0.75 | — | — |

* Previous results from Table 2 for no frame are reported for better comparison without BioTF representing.

Table 8

The comparison accuracy of the traditional and time frame signal processing using high-low series. The best results in this experiment are in bold and all experiments are in italics.

| Data | Train (Test) Accuracy /Test F1 score over frame length l | | | | | |
|------|--|-------------------------|-------------------------|-------------------------|-----------------------|-------------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 81.9 (81.2)/ 0.81 | 80.5 (80.3)/ 0.79 | 79.9 (80.0)/ 0.79 | 80.2 (80.1)/ 0.79 | 80 (79.9)/ 0.78 | 79.3 (78.5)/ 0.77 |
| | | | | | | |
| | | | | | | |
| EEG | 83.3 (83)/ 0.825 | 86 (85.9)/ 0.85 | 83.4 (83.4)/ 0.82 | 81.1 (81.2)/ 0.80 | — — | — — |
| | | | | | | |
| | | | | | | |
| ECG | 80.5 (82)/ 0.715 | 89.5 (71.1)/ 0.70 | 86.2 (71.1)/ 0.71 | 75.3 (70.4)/ 0.68 | — — | — — |
| | | | | | | |
| | | | | | | |

* Previous results from Table 2 for no frame are reported for better comparison without BioTF representing.

batch normalization layer. The layer normalizes a mini-batch of data across each channel for each independent observation. Table 5 shows the performance results. Results show that the accuracy of all signals increases using the time frame approach.

- Adding dropout layer

A dropout layer after the LSTM layer is used to prevent network overfitting with a dropout probability of 0.2 that randomly sets the input to zero. This layer is added to EMG and ECG networks. The EEG network already has a dropout layer. Table 6 shows the performance results. Results show that the accuracy on all signals increases using the time frame of 5.

3.4.3. Varying input dimension

We change the input dimension by adding or removing a variate in this part.

- Adding signal energy dimension

We first add the signal energy dimension. For energy calculation, we use the Fourier transform as follows:

$$\text{Signalpower} = \int |\mathcal{X}(f)|^2 df, \quad (5)$$

where $\mathcal{X}(f)$ is the Fourier transform of $x(t)$. Therefore, the dimension increases to five for each channel. Table 7 shows the performance results. Results show that the accuracy in EMG signal increases using the time frame of 5.

Table 9

The comparison accuracy of the traditional and time frame signal processing using high-low difference series. The best results in this experiment are in bold and all experiments are in italics.

| Data | Train (Test) Accuracy /Test F1 score over frame length l | | | | | |
|------|--|---------------------------------|-------------------------|-------------------------|------------------------|-------------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 81.1 (81.2)/ 0.81 | 81.4 (81.6)/ 0.80 | 80.7 (80.2)/ 0.79 | 79.8 (79.6)/ 0.78 | 79.9 (79.6)/ 0.7 | 79.2 (79.3)/ 0.78 |
| | | | | | | |
| | | | | | | |
| EEG | 82.9 (83)/ 0.825 | 66.9 (66.3)/ 0.59 | 70.9 (70.6)/ 0.66 | 73.4 (72.3)/ 0.68 | — — | — — |
| | | | | | | |
| | | | | | | |
| ECG | 80.5 (82)/ 0.71 | 92.0 (86.5)/ 0.865 | 77.4 (62.8)/ 0.62 | 84.8 (82.8)/ 0.82 | — — | — — |
| | | | | | | |
| | | | | | | |

* Previous results from Table 2 for no frame are reported for better comparison without BioTF representing.

Table 10

The comparison accuracy under pruning architectures. The best results in this experiment are in bold and all experiments are in italics.

| Data | Train (Test) Accuracy /Test F1 score over frame length l | | | | | |
|------|--|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 51.9 (51.8)/ 0.45 | 81.7 (81.8)/ 0.81 | 81.9 (81.7)/ 0.79 | 81.5 (81.6)/ 0.79 | 81.2 (81.0)/ 0.78 | 79.8 (79.0)/ 0.76 |
| | | | | | | |
| | | | | | | |
| EEG | 79.1 (78.7)/ 0.78 | 79.4 (79.5)/ 0.79 | 81.9 (81.8)/ 0.79 | 77.4 (77.7)/ 0.73 | — — | — — |
| | | | | | | |
| | | | | | | |
| ECG | 68.1 (71.2)/ 0.70 | 53.8 (50.7)/ 0.50 | 50.8 (51.0)/ 0.45 | 52.8 (46.9)/ 0.45 | — — | — — |
| | | | | | | |
| | | | | | | |

* Means no frame has been applied, and the original signal has been processed without BioTF representation.

- Using high and low series

Then, we decrease dimensions by removing starting and ending characteristics and only considering high and low. Therefore, we have a bivariate signal to process for each channel. Table 8 shows the results. Results show that the accuracy of EEG signals increases using the time frame $l = 5$.

- Using high-low difference series

Finally, we use only the difference between the high and low series to generate a univariate time series from each channel. Results (Table 9) show that the accuracy on all signals increases using a time frame $l = 5$ except EEG. The classification accuracy on the ECG dataset reached the highest value of 86.5 percent in all experiments.

3.4.4. Pruning architectures

In this part, we perform the same experiments with different architectures. Specifically, we decrease the LSTM units by a tenth for EMG and ECG datasets. Also, we only keep the last block (each block ends with an Identity layer) in EEG signal network architecture. These experiments aim to find how the data representation could lead to simple architectures and the possibility of real-time implementation [13]. Table 10 shows the accuracy results. Results show that the accuracy of all signals increases using the time frame approach except for ECG. This pruning is also beneficial for neuroimaging processing by CNNs [43] and muscle fatigue tracking [44] because the learning parameters could significantly decrease. So, the model could be trained on smaller datasets and online.

3.4.5. Changing frame slide

In all previous experiments, time frames have no overlap, while biosignals such as EMG signal features are often extracted from overlapped window segments. In this section, we use time-frame $l = 50$ with a 10-sample overlap, time-frame $l = 25$ with a 5-sample overlap, and time-frame $l = 10$ with a 2-sample overlap for the EMG signal. Also, we use a high-low difference of time-frames $l = 5$ with a 2-sample overlap for the ECG signal. The test $l = 5$ obtains %89.61 for the EMG signal. Also, for the ECG signal, accuracy hits the accuracy of 92.98 %. For both datasets, changing the frame slide achieves the highest accuracy.

3.4.6. Changing window size

All the previous experiments use the same window size per dataset. Considering all previous experiments use a window size of n , here, we examine the effect of window sizes $n/2$ and $n/4$ on traditional and BioTF biosignal processing. Therefore, the training and testing datasets become double and fourfold. Note that the pruned architecture of the convolutional network for EEG is used due to the small feature map of window sizes of $n/2$ and $n/4$. Table 11 shows the accuracy results drop

Table 11

The comparison accuracy under various window sizes. The best results in this experiment are in bold and all experiments are in italics.

| Data | Test Accuracy over frame length l using window sizes of $n, n/2, n/4$. | | | | | | | | |
|--|---|-----------------|-----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|------------------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ | | | |
| EMG | 81.2, 82.82, 82.7 | 80,79,9,79,6 | 79,4,79,1,79,7 | 79,8,78,6,78,8 | 79,2,78,2,77,4 | 79,77,5,75,7 | | | |
| EEG | 83,82,5,81,7 | 82,82,34,83,51 | 84,80,63,82,97 | 83,-,- | - | - | | | |
| ECG | 82,64,8,74,3 | 70,64,8,73,6 | 55,63,3,75,2 | 72,62,0,63,9 | - | - | | | |
| Test accuracy of Maintenance task on original signal over various feature sets (average on cross-validations and sessions) using window sizes of $n, n/2, n/4$. | | | | | | | | | |
| | D1 [#] | D2 [#] | D3 [#] | D4 [#] | D5 [#] | D6 [#] | D7 [#] | D8 [#] | F1 ^{\$} |
| HD-sEMG [†] | 93,8,89,5,82,7 | 91,1,84,5,76,2 | 92,4,86,9,80,4 | 92,3,88,2,81,4 | 93,7,88,4,80,4 | 87,3,80,2,64,6 | 90,80,6,68,3 | 90,2,86,0,75,3 | 93,5,93,1,93,1 |

* Means no frame has been applied, and the original signal has been processed without BioTF representing. [#]Data point-based feature set. ^{\$} time frame-based feature set. [†]Only original signals of maintenance tasks are included from the HD-sEMG dataset.

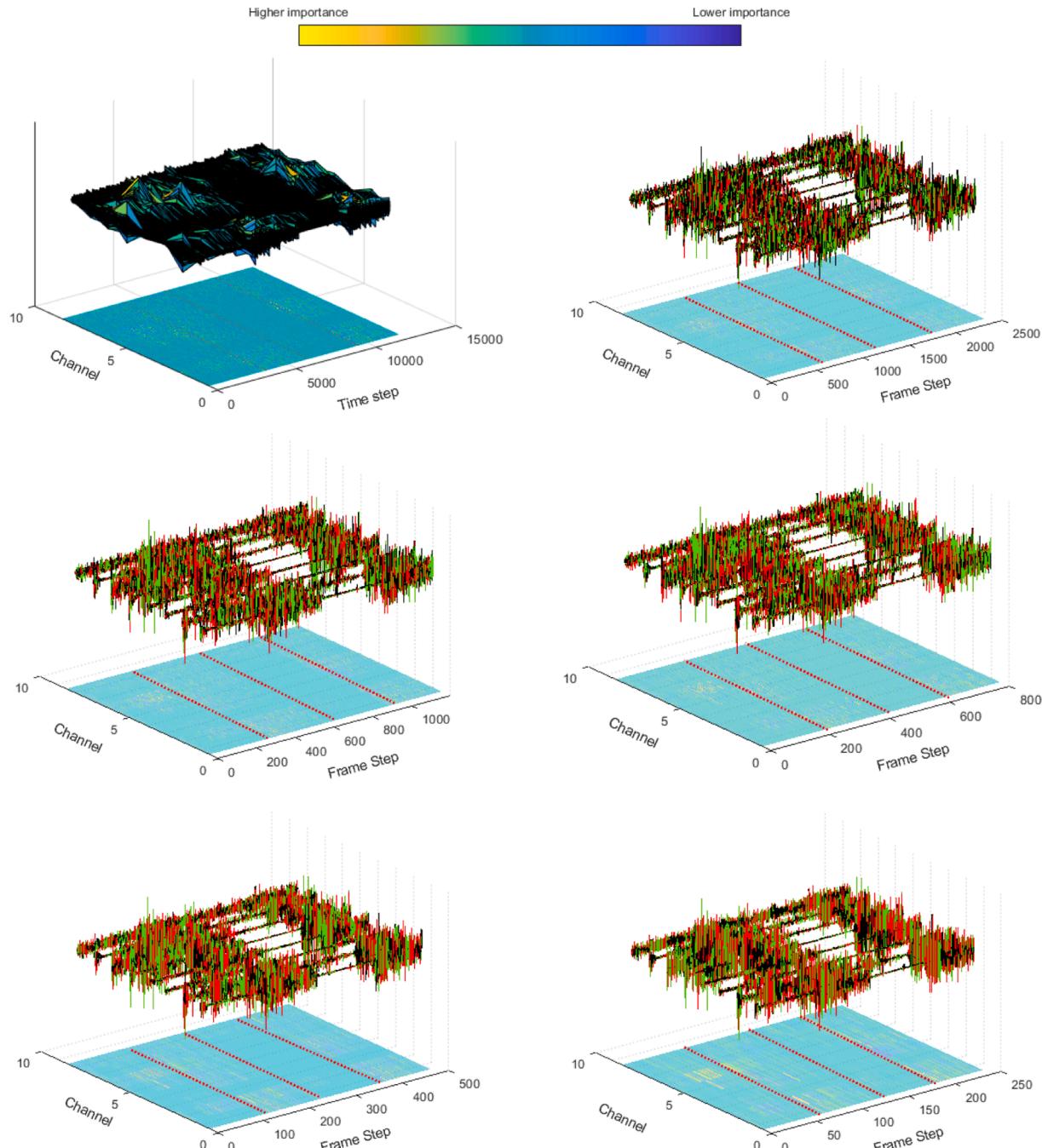


Fig. 5. Explanation of EMG for HandOpen, WristFlexion, Rest, and WristFlexion gestures at time frames $l = 1, l = 5, l = 10, l = 15, l = 25$, and $l = 50$ from left to right. The average importance of characteristics is shown for candle importance in time frame-based processing. Magnified views are provided in Fig. 6.

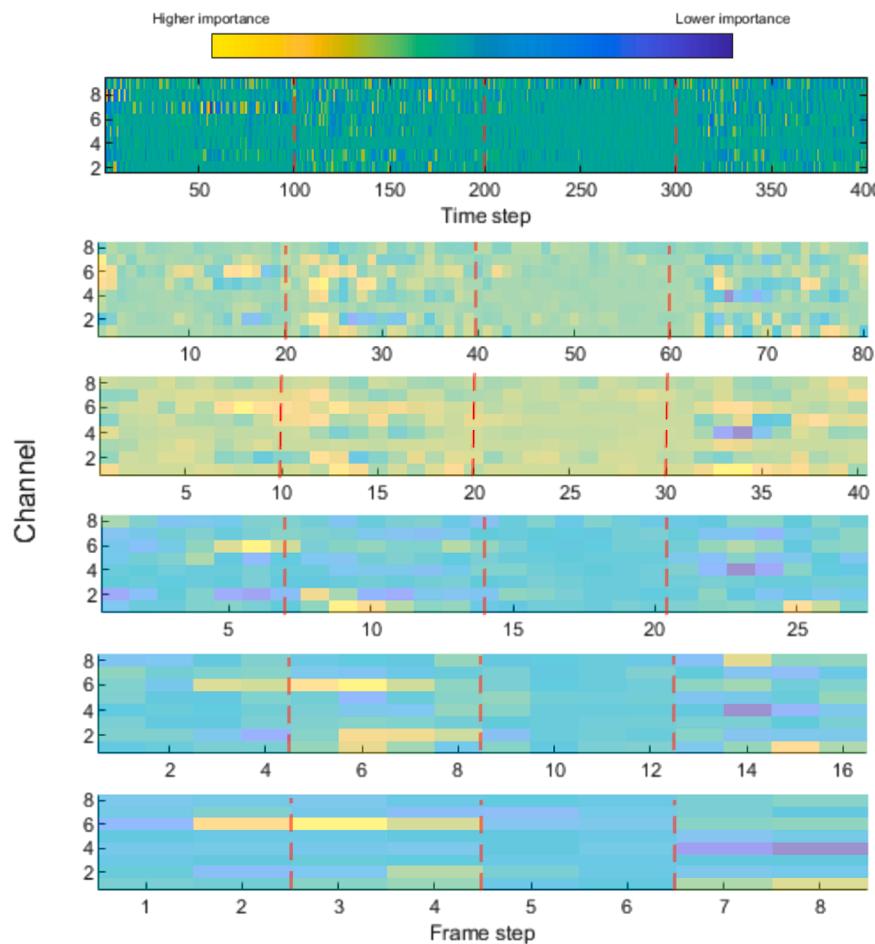


Fig. 6. Magnified frame length of $l = 1, l = 5, l = 10, l = 15, l = 25$, and $l = 50$ from top to bottom. Vertical dash lines separate HandOpen, WristFlexion, Rest, and WristFlexion gestures.

in most cases when we decrease window size for both methods.

3.5. Explainability results

The explainability of the decision-making system plays an important role in sustainable and practical usage. Here, we provide explainability results for all datasets except HD-sEMG due to its classification model limitations. We compare the explanation results of time frame representation against the original signal over randomly selected samples. The classification explanation of these samples for both approaches is then performed for better comparisons. Here, the LIME method for the EEG network and the activation of the input layer for other networks over traditional data points and time frame candlesticks ($\bar{x}_k(i)$) are calculated. More details on the parameters of the explanation methods are provided in Appendix I.

Figs. 5-12 shows the signal part contribution to the predicted class. The provided results on EMG (Fig. 5) show that channels 1, 3–6, and 7 in all experiments are important to determine related gestures that are consistent with previous works. Specifically, channel 6 is important for the HandOpen gesture, and channels 1,3–5, and 7 are related to HandFlexion [45]. To provide more details, we first apply a moving average

with a window size of 50 and then decrease the length by a down sampling of rate 30. Fig. 6 shows the scaled results of each experiment for the same sample as Fig. 5. The important regions in all experiments are similar, but as we consider longer frames, the importance becomes easier to follow.

The explanation for EEG (Figs. 7-9) for all networks shows classification dependence on the same temporal data. However, an explanation of the proposed time frame processing signal reveals regions (spatial) consistent with the highly related channels to determine the normality/abnormality [46]. In other words, time frame processing provides a temporospatial explanation over EEG signals (compare the yellow regions). Also, these temporospatial data are almost the same for time frames. Overall, the networks focus on occipital region recordings in time frame processing for abnormality detection. Fig. 9 shows only the magnified explanation for BioTF classifications on regions of interest because of the visualization limitation of the datapoint-based network. As l increases, the region of interest becomes longer.

All the networks mainly rely on the extremum points or pivot time frames because peak values and durations are essential features in ECG signals (as shown in Figs. 10-12) [47]. To better focus on the differences, we plot the results for the first 750 abnormal and normal ECG data

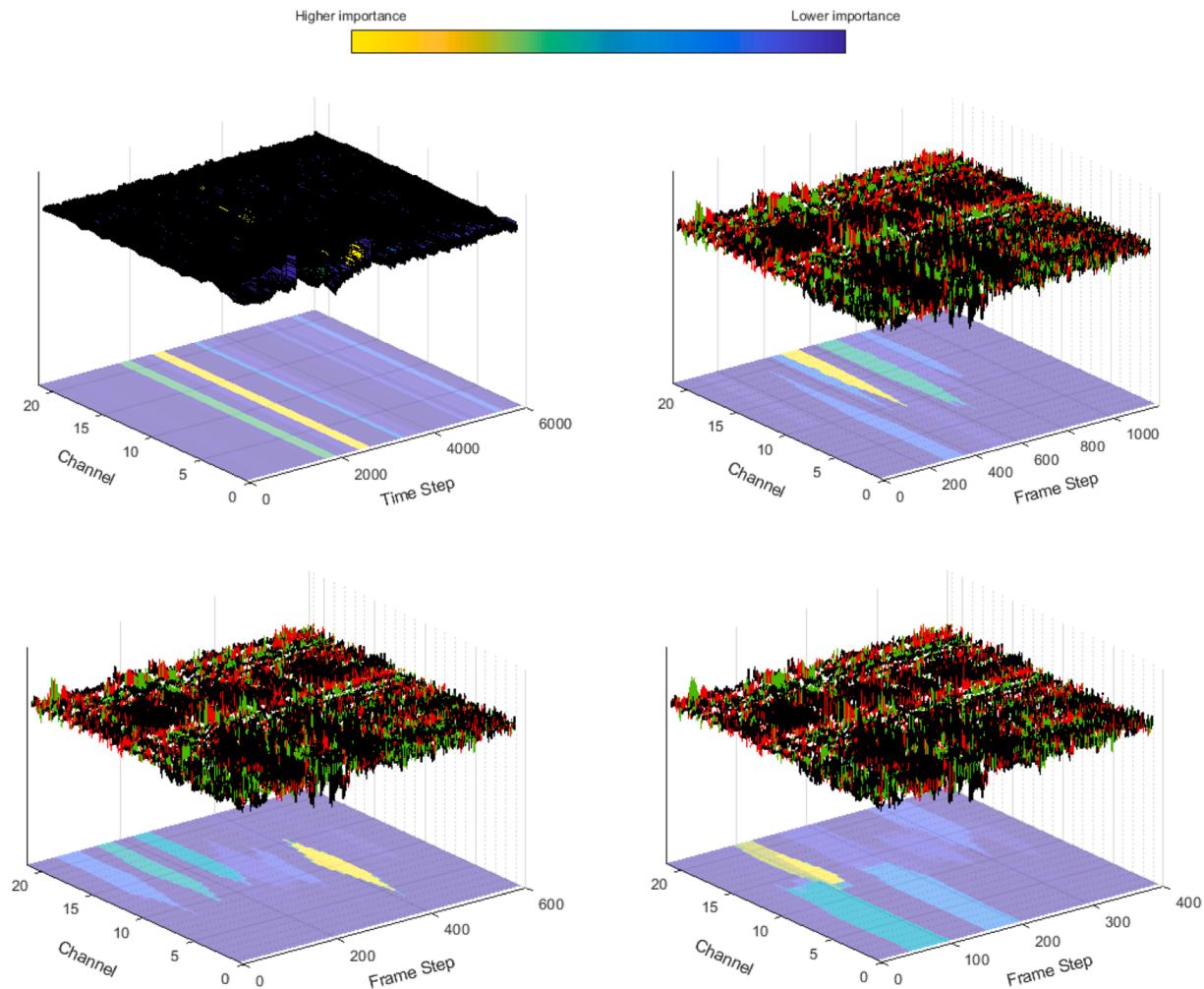


Fig. 7. Explanation of abnormal EEG at time frames $l = 1$, $l = 5$, $l = 10$, and $l = 15$ from left to right. The average importance of characteristics is shown for candle importance in time frame-based processing.

points. As shown in Fig. 12, the data point-based and frame-based processing results are consistent. However, networks focus on a time spectrum in data point-based processing, while in time frame-based processing, specific characteristics are a matter of concern, similar to physicians' analysis.

3.6. Computational time

We compare the training computation time of architectures, which is a significant factor in biological signal processing [48]. This significance arises from biosignals' time-series nature and complexity, often requiring sophisticated architectures that can lead to extended training times. The computational demands can become particularly notable when employing deep learning and end-to-end processing approaches.

Table 12 shows the average training time of networks. As shown, computation time significantly decreases when the sequence length is long, such as in EMG, EEG, and ECG. The sequence length of the EEG, EMG, and ECG signals, decreased by a scale of 1/5, 1/10, 1/15, 1/25, and 1/50 using frame lengths of 5, 10, 15, 25, and 50, respectively, overcomes the increased computation due to the fourfold increased dimension by frame characteristics. For the HD-sEMG dataset, we do not

measure the computational time because the processes involve separate feature extraction and classification stages. However, it is evident that features like frame body length (FBL) offer signal compression, leading to less memory and computation than data point-based ZC and sign slope change (SSC) feature counterparts [37].

4. Discussion

The proposed BioTF method enhances classification accuracy, as shown in Sections III.C&D. It performs better in datasets with high noise levels but less so in datasets with low noise levels, such as ECG. For the latter case, a lower correlation achieved by using fewer frame components or applying differentiation could yield better results (see Section III.D.c).

The proposed time frame processing also provides a visualization that shows the temporospatial significance of bio series (except for mono channel ECG signal) in contrast to the temporal dependencies of regular processing (see Figs. 5-9). This temporospatial information could help identify the critical channels (sensors) and regions that influence the networks' final decision. Thus, by examining the visualization, one could choose a subset of sensors for decision-making instead of using all

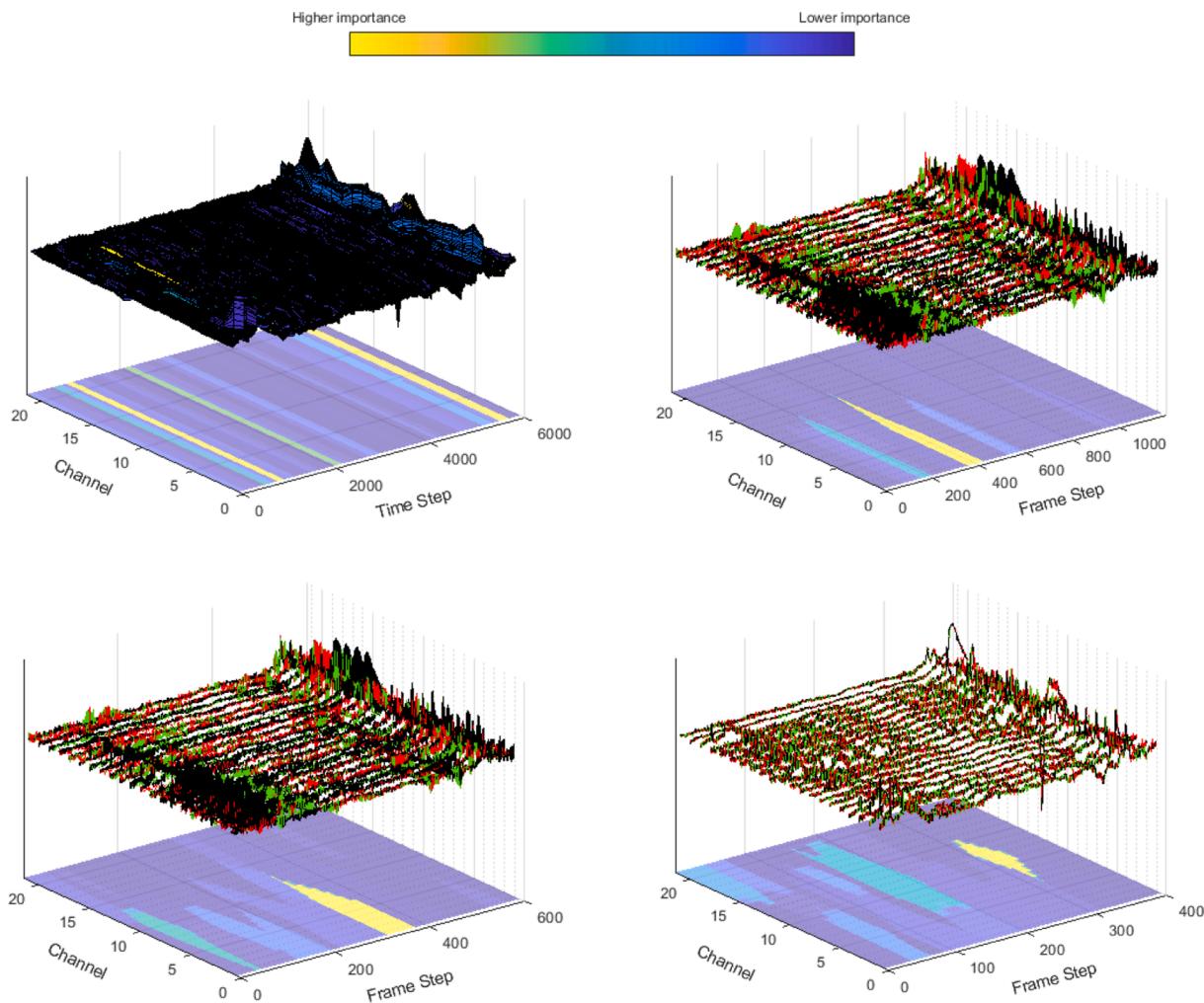


Fig. 8. Explanation of normal EEG at time frames $l = 5$, $l = 10$, and $l = 15$ from left to right. The average importance of characteristics is shown for candle importance in time frame-based processing.

of them, which could reduce costs considerably.

Besides, the training time dramatically decreases by the proposed time frame-based processing (see Table 12). The processing idea is consistent with the promising sequence-aligned recurrence approaches [49], aligning SHLE values of frames. This reduction in training computation time allows us to conduct various machine learning experiments quickly and mitigate memory constraint issues, especially for long sequence lengths in cloud computing environments. Also, models with lower computation costs, such as the proposed processing approach, are suggested for clinical translations [37,50]. Further studies could be directed to investigate the performance of BioTF under a trimmed dataset, decreasing data collection efforts.

Also, one of the main challenges for deep learning algorithms is model complexity, which limits hardware implementation [50], even when we want to fine-tune the network with new data [51] since complex models require proper parameter tuning [52] and large datasets to train [43]. The proposed processing achieves better results with simpler pruned architectures (see Section III.D.d) and reduced learning

parameters. Thus, it is more suitable for real-time learning, such as processing gadgets' hardware and clinical settings.

The accuracy of the ECG dataset shows high variations, while other datasets are more stable. Also, the traditional signal processing outperforms the proposed method with basic components in the ECG signal. The time frames for ECG signals are $l = 5$, $l = 10$, and $l = 15$ here. Considering the 300 Hz sampling rate, time frames take about 0.016, 0.033, and 0.05 s. However, the highest frequency band needs 0.01 (Table 1) to repeat. The time frame representation above this time length may obscure valuable information. Therefore, we present experiments with overlapped time frames in appendices with promising results (Section III.D.e).

Several CNN/LSTM architectures sequentially process spatial/temporal patterns. However, extracting spatiotemporal patterns from them is difficult. As stated, LAXCAT [14] incorporates spatial and time interval information to focus on the discriminative pattern. Also, XAI4EEG [15] recently identified effective spectral, spatial, and temporal regions in EEG signals. This method is demonstrated for EEG signals here;

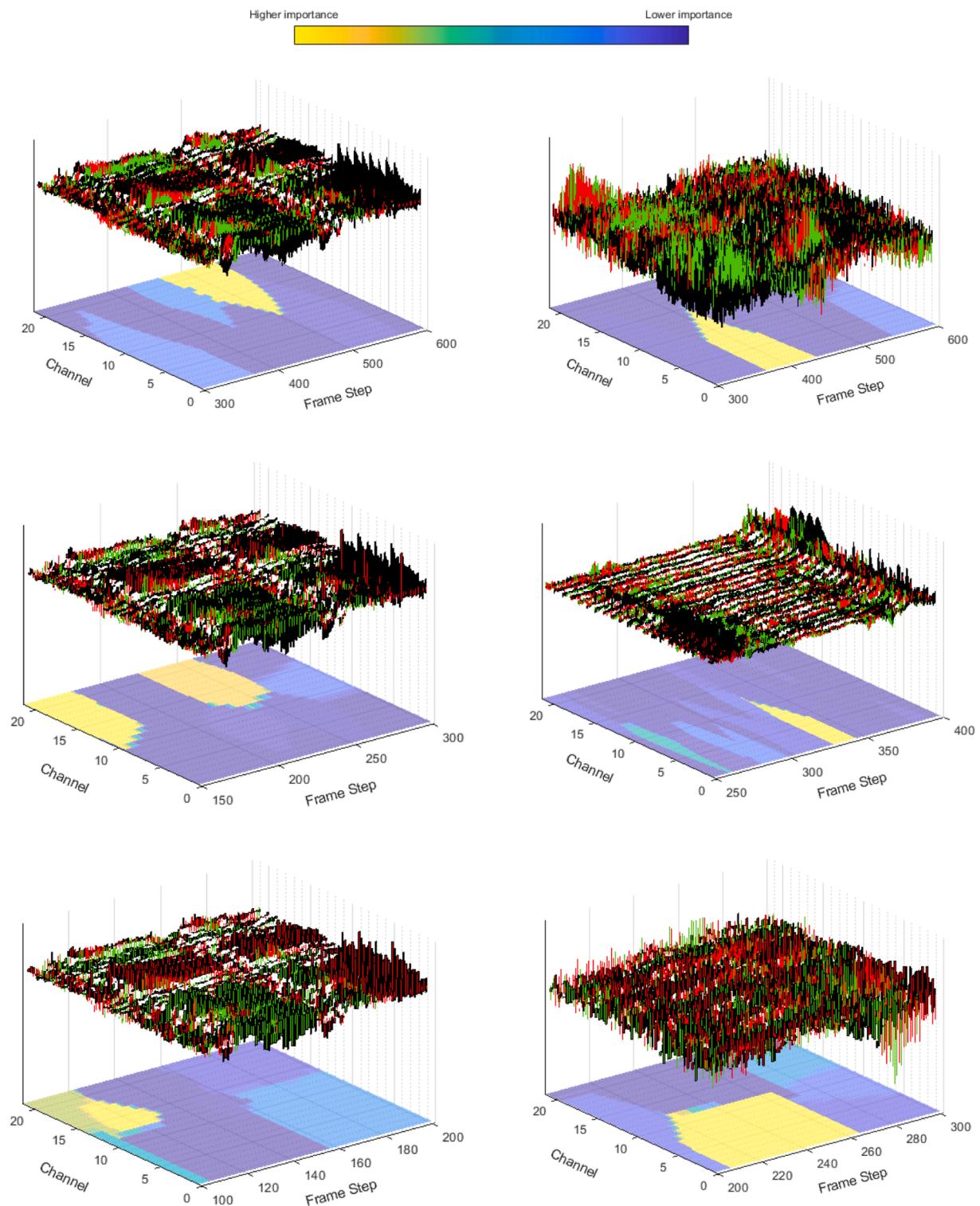


Fig. 9. Explanation (magnified) of abnormal (left column) and normal (right column) EEG at time frames $l = 1$, $l = 5$, $l = 10$, and $l = 15$ from up to down. The average importance of characteristics is shown for candle importance in time frame-based processing.

however, the proposed method applies to any architecture. As Figs. 7–9 illustrate that the proposed time frame representation reveals the temporospatial explanation within the EEG signal, unlike traditional processing that only emphasizes temporal features.

Another reason behind the time frame processing financial time series is the unevenly spaced price releases, known as tick data. Although current bio datasets are often recorded in time-even modes, the proposed BioTF provides a reasonable basis for uneven biosignal

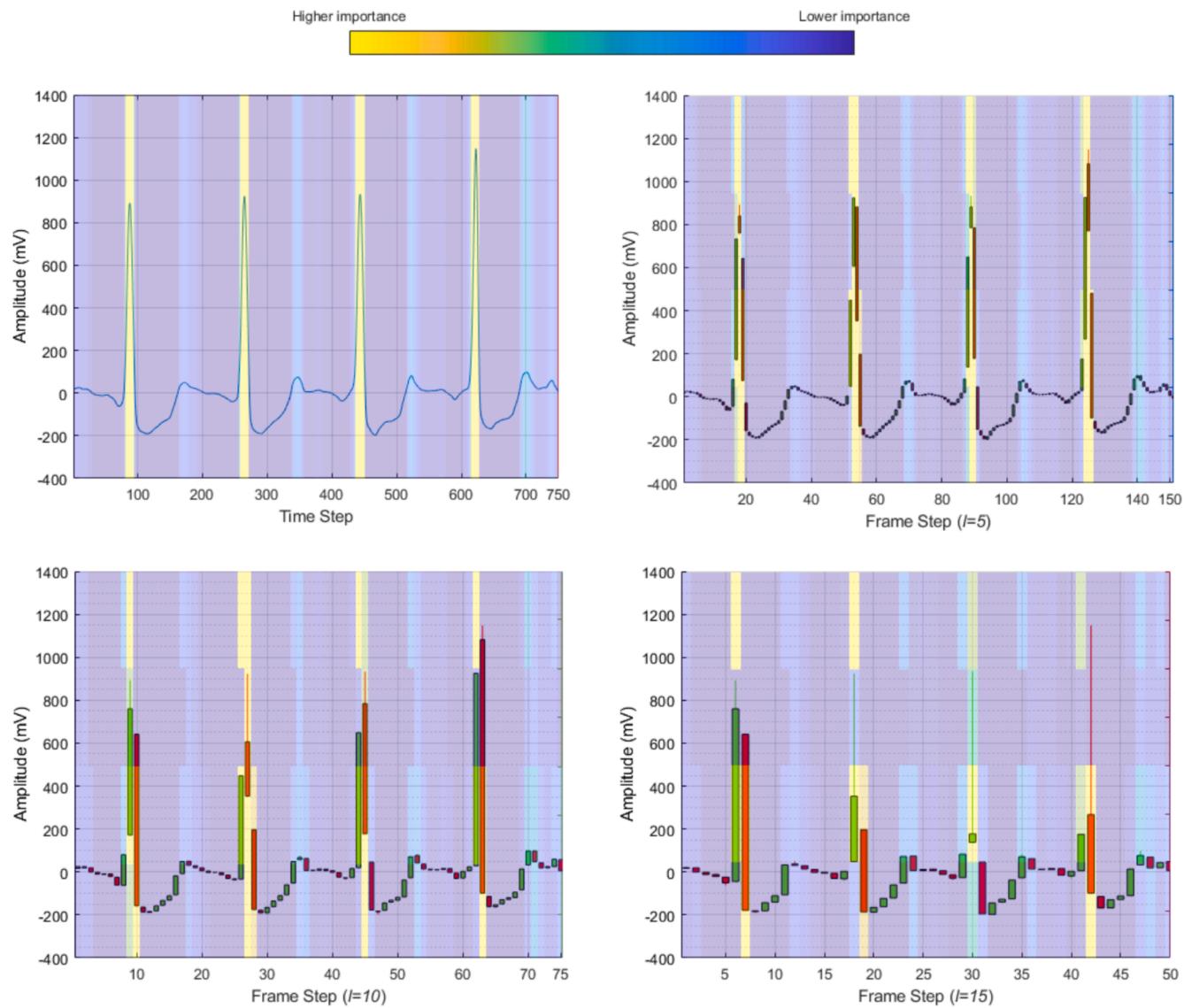


Fig. 10. Explanation of abnormal ECG at time frames $l = 1$, $l = 5$, $l = 10$, and $l = 15$ from left to right on specific steps for higher temporal resolution. The average importance of characteristics is shown for candle importance in time frame-based processing. More views are provided in Fig. 12.

processing. This bio-processing can be helpful for devices that record biosignals for the long term with minimal memory and power usage. Furthermore, these recordings are often affected by noises, disturbances, and motion artifacts, which make the frames more meaningful.

Several limitations must be considered. First, if we aim to create rigid time frames, there may be a lag when calculating components in real-time for specific time frames. An alternative approach is to utilize flexible shape candlesticks that can be dynamically updated at any data point. Second, the current representation characteristics do not adequately capture the signal distribution within longer time frames, which is a crucial consideration.

Finally, the proposed BioTF is a general framework, and any machine

learning and deep learning with any degree of complexity, such as [16–19,53], and other biosignals [54–56] can use BioTF representation instead to improve efficiency.

The two general remarks below address the above results from the perspective of potential medical users and the approach's advantages as a classifier.

Remark 1. *The main representation results are summarized as follows:*

- *The time frame representation of biosignals is intuitive and similar to what humans perform, observing the extrema of sequential frames. For example, physicians calculate wave characteristics based on sheet squares in ECG signals.*

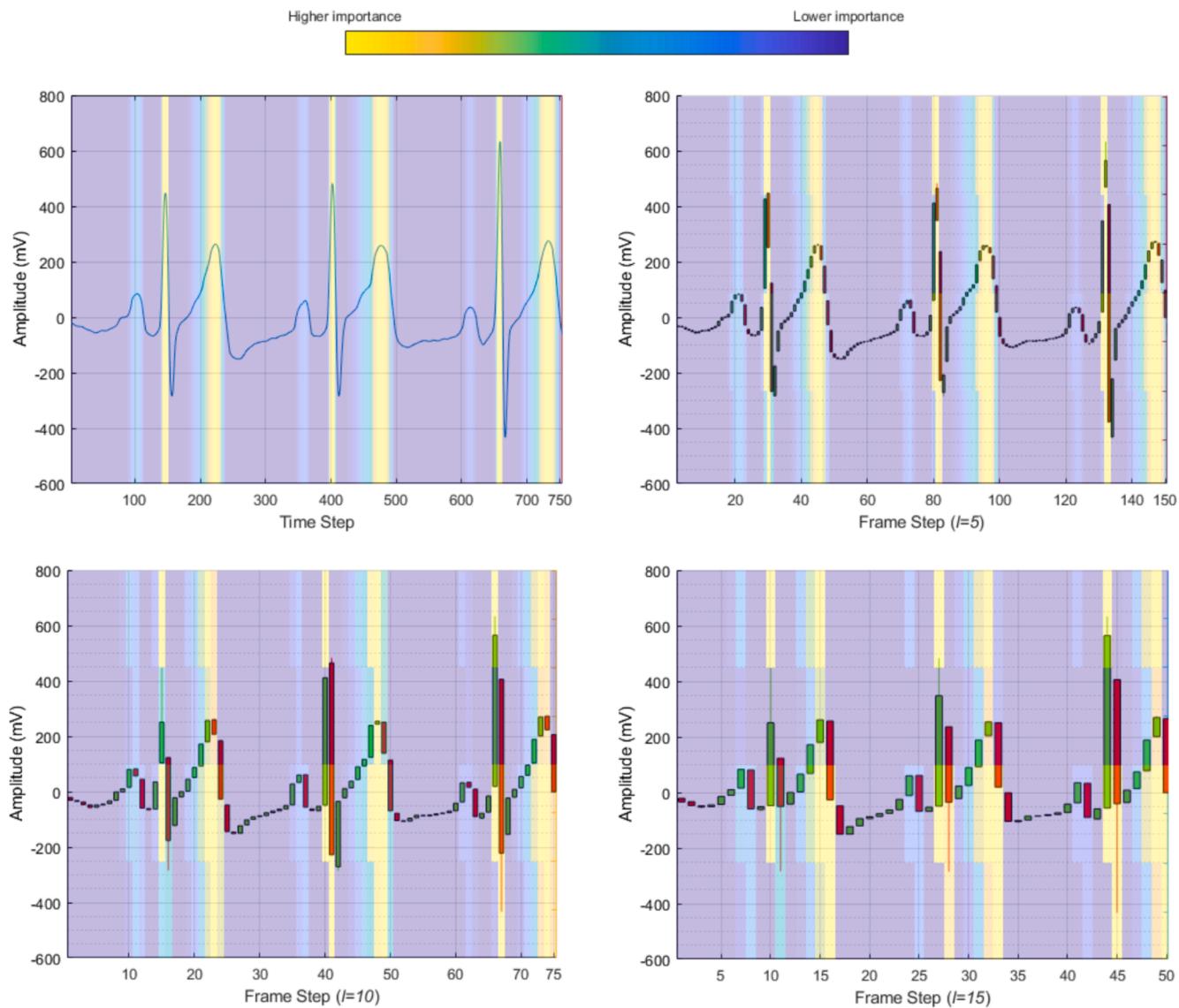


Fig. 11. Explanation of normal ECG at time frames $l = 1$, $l = 5$, $l = 10$, and $l = 15$ from left to right on specific steps for higher temporal resolution. The average importance of characteristics is shown for candle importance in time frame-based processing. More views are provided in Fig. 12.

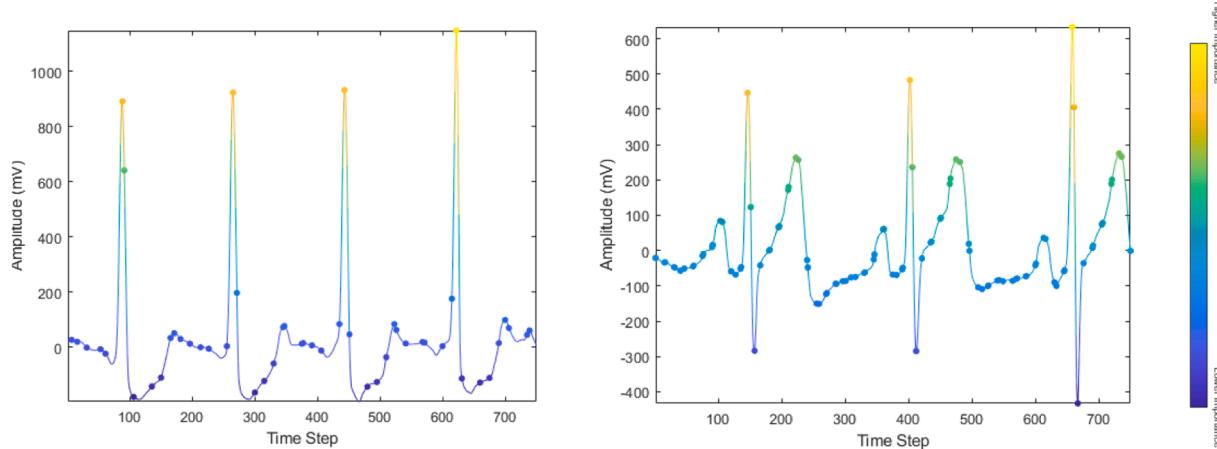


Fig. 12. The comparison of data-point (colorful line) and frame-based (colorful dots) processing for abnormal (left) and normal (right). Here, only the highest values within a frame ($l = 15$) are shown. Please note that the number of colorful dots is equal to the number of frames which is 50 for 750 data points.

Table 12

Computational times (in minutes) of experiments.

| Dataset | Computation time over frame length l | | | | | |
|---------|--|---------|----------|----------|----------|----------|
| | $l = 1^*$ | $l = 5$ | $l = 10$ | $l = 15$ | $l = 25$ | $l = 50$ |
| EMG | 198 | 47 | 27 | 20 | 14 | 11 |
| EEG | 93 | 46 | 33 | 22 | — | — |
| ECG | 2663 | 507 | 214 | 176 | — | — |

* Means no frame has been applied, and the original signal has been processed without BioTF representing.

- They visually explain the situation of increasing or decreasing a biosignal. Physicians could determine the state of increasing/decreasing by just looking at the color and frame length.
- They help detect abnormality and latent patterns through a simplified, understandable visualization.
- The proposed time frame representation is a generalized approach to extracting features from biosignal partitions by extending feature space exploration to three dimensions (amplitude, time, and frame). If $l = 1$, this processing reverts to data point-based.
- The proposed BioTF encourages the adoption of further financial analysis for biosignal processing and repurposing financial indicators to quantify biosignal excitements.

Remark 2. Several advantages could be extracted from the above classification results as follows:

- The proposed representation offers competitive accuracy as the data point-based processing (see Table 2-10); however, it could be superior by mixing feature sets for HD-sEMG and using component combinations instead of just SHLE components for the rest of them.
- The proposed BioTF improves explanations over the spatial and temporal domains, especially in EEG signals.
- The networks' performance drops because of components with high correlations for smooth biosignals such as ECG (see Table 2). In this case, the combination of candlestick components could improve performance.
- The time frame length l significantly affects accuracy (see Table 2).
- The proposed time frame-based processing makes training faster (see Table 12).
- The proposed method could reduce memory usage, which is important in high dimensional signal acquisition and processing tasks, such as HD-SEMG [53].
- The proposed processing leads to simpler architectures. Therefore, it is more proper for real-time processing of gadgets hardware (see Section III. D.d).
- Also, the proposed BioTF provides a basis for low-power, long-lasting bio recordings using gadgets.

5. Conclusion

Biosignal processing algorithms often work on highly volatile and noisy data points, leading to challenges in computation and explanation. Here, based on the similarity between biological and financial time series, we propose to use data frames as a processing unit. Specifically, we propose to extract starting, highest, lowest, and ending values within a frame and transfer a series of data points into the data frames. This series of frames forms the basis for further classification and post hoc understandable explanations. Results show that the proposed BioTF provides

competitive classification in the HD-sEMG dataset and around 1 % improved accuracy on the average in the remaining datasets. Also, the proposed time frame representation better explains the patterns compared to the original signal by displaying the signal's temporospatial importance.

Furthermore, training computation time significantly decreases using the proposed BioTF, enabling various machine learning experiments with fewer computations and, hence, feasible clinical implementations. This approach offers several key advantages: it provides a more intuitive representation of complex biosignals, leverages cross-domain knowledge from financial analysis, enhances the AI decisions' understandability, and improves computational efficiency. These properties contribute to creating AI systems that are not only accurate but also more transparent, efficient, and potentially more accessible for healthcare professionals to understand and use. Future work could further explore the potential of historical financial patterns and technical analysis for more transparent biosignal classifications. Further studies emphasizing the specific applications of HD-EMG are also recommended.

Code availability

The implementation can be accessed at <https://github.com/Hamed-Rafiei/BioTF>.

Data availability

No data was used for the research described in the article.

Funding

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Hamed Rafiei: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mohammad-R. Akbarzadeh-T:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A: A closer look at the time frame representation of biosignals

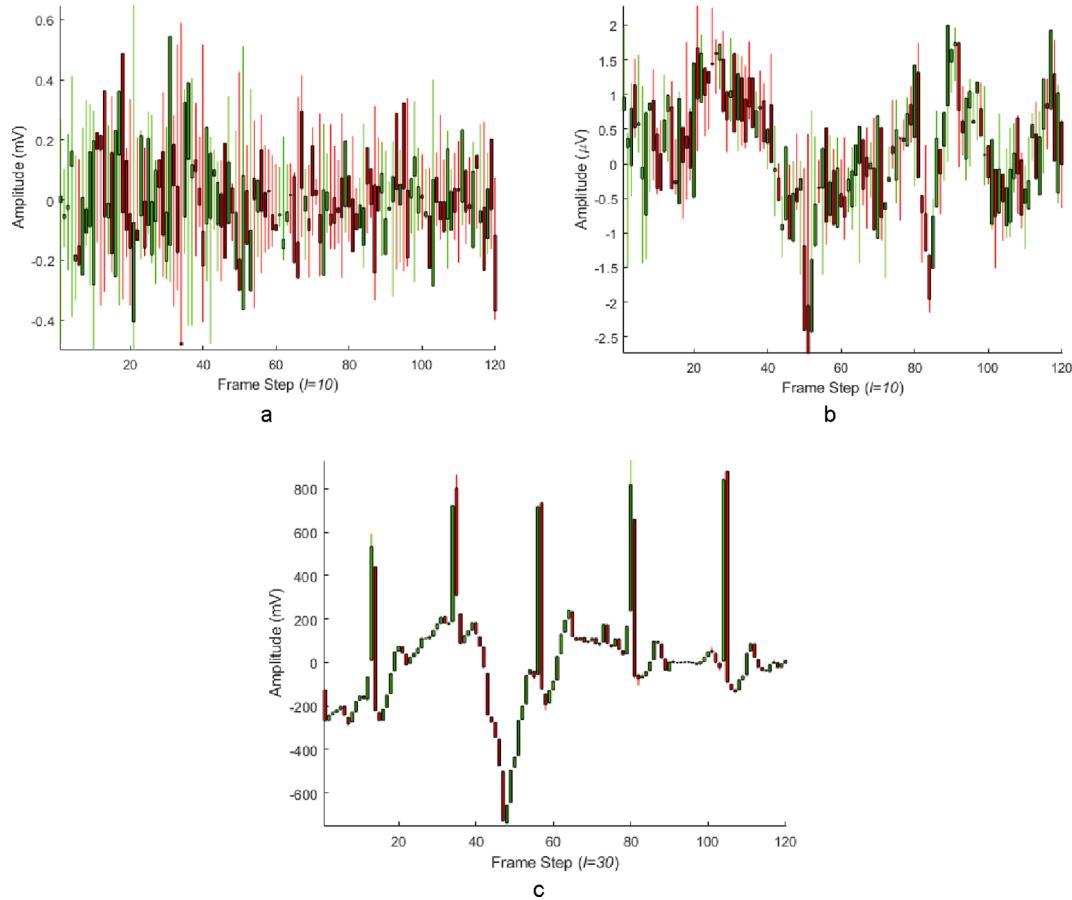


Fig. A1. Candlestick frame representation of a) EMG, b) EEG at , and c) ECG at .

Fig. A.1 shows the sample time frame representation of biosignals.

Appendix B: Feasibility of time frame classification

Fig. B.1 shows the first channel of EMG signal recorded from forearm muscles (sequentially performing WristsFlexion, HandOpen, Rest, and WristExtension gestures) for the Rickshaw Man pattern. The starting and ending values in the Rickshaw Man pattern are very close together with long shadows. This pattern shows that a specific change in signal is not expected. A time frame $l = 200$ is used in the illustration. As shown, this pattern appears when the EMG signal rests. Based on the results mentioned, one could conclude there is potential for finding specific patterns for each class.

Technical analyses provide insight into the market. The works closest to these analyses are provided by [32] for grasp force estimation and [33,34] for determining the contraction onset using the MAV of the EMG transient signal. Fig. B.2 shows the results of applying *atr* indicators on the EMG biosignal. Results show that technical analysis has a good potential for discriminating the EMG classes, having different levels for each class.

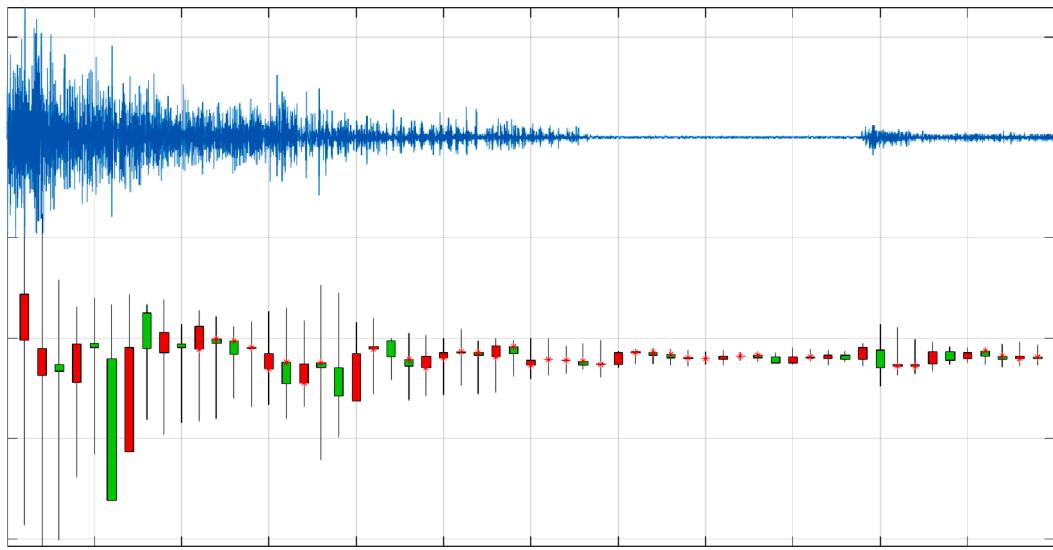


Fig. B1. Applying neutral financial pattern on EMG signal.

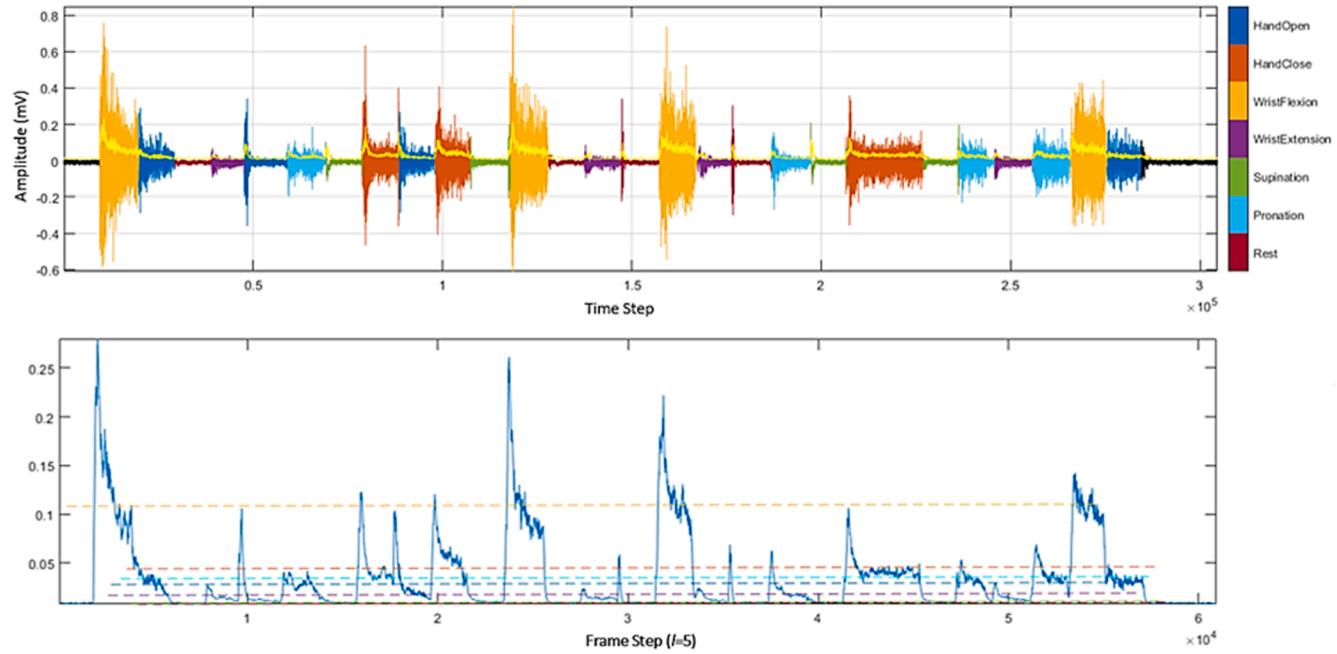


Fig. B2. The upper plot shows the EMG signal from the first channel for various gestures (the yellow line shows a simple moving average with window 100 over absolute values). The bottom plot shows the atr technical indicator for the smoothed values (simple moving average with window 10) of time frame with period 100 and related colorful levels of gestures.

Appendix C.: Technical analysis

In this part, we calculate different financial technical analyses to observe the signal's property, direction, and volume. This analysis is provided for each class of the EMG dataset. Here, we analyze the SHLE components of the time frame $l = 5$ to have minimal latency. Because we do not modify the technical analysis formula, all the values are added by constant numbers without losing generality to prevent negative values (prices are not negative). Oscillators often show the momentum of extreme situations. Stochastics compare the ending characteristics of a time frame with past time frames. Indexes show signal movement speed and change. Fig. C.1 shows the results of applying oscillator, stochastic, and indicator operators on EMG biosignal.

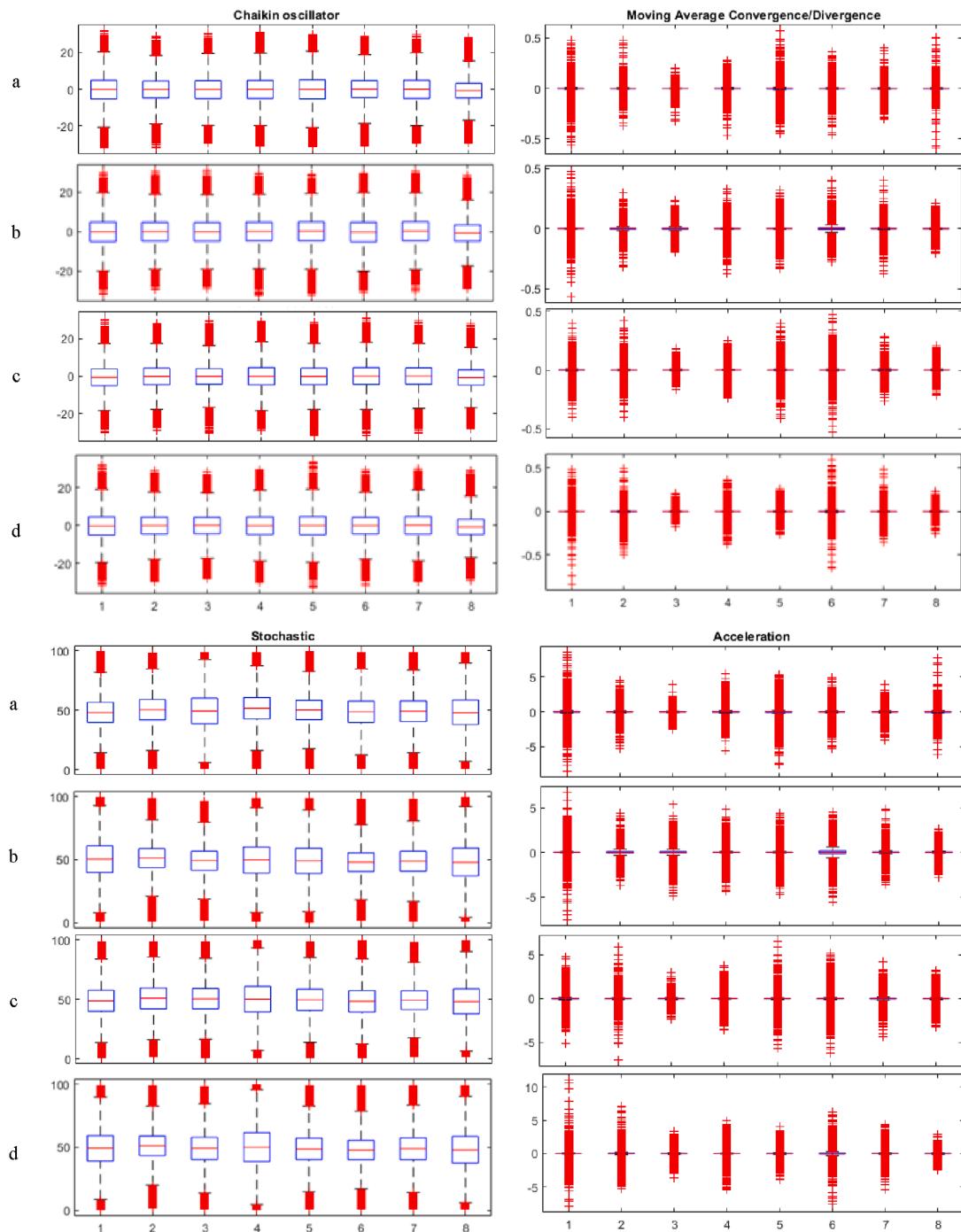


Fig. C1. Technical analysis of EMG signal using financial indicators over eight channels for a) wrist flexion, b) wrist extension, c) hand close, and d) hand open gestures. The last two rows are related to Lowest Low indicator.

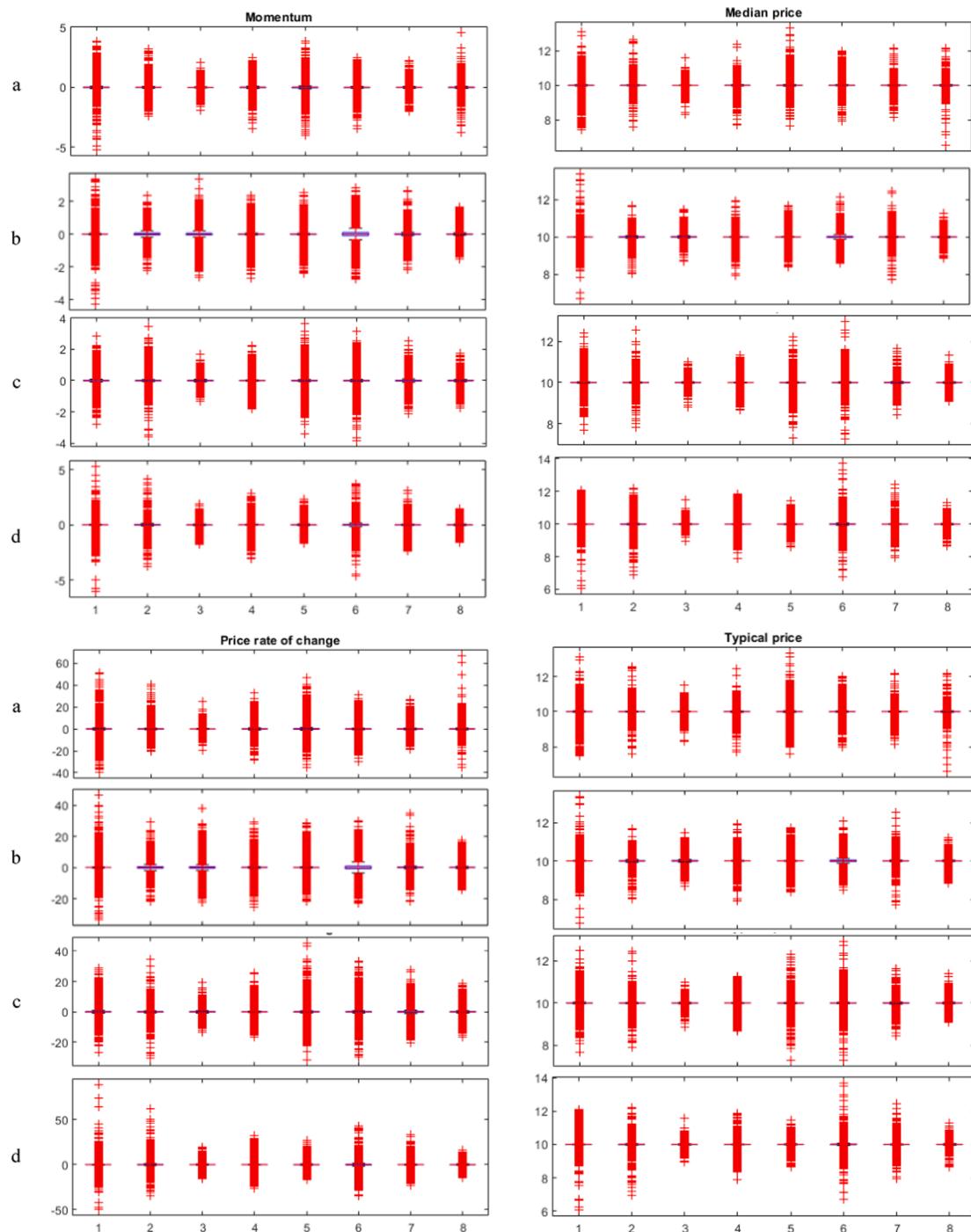


Fig. C1. (continued).

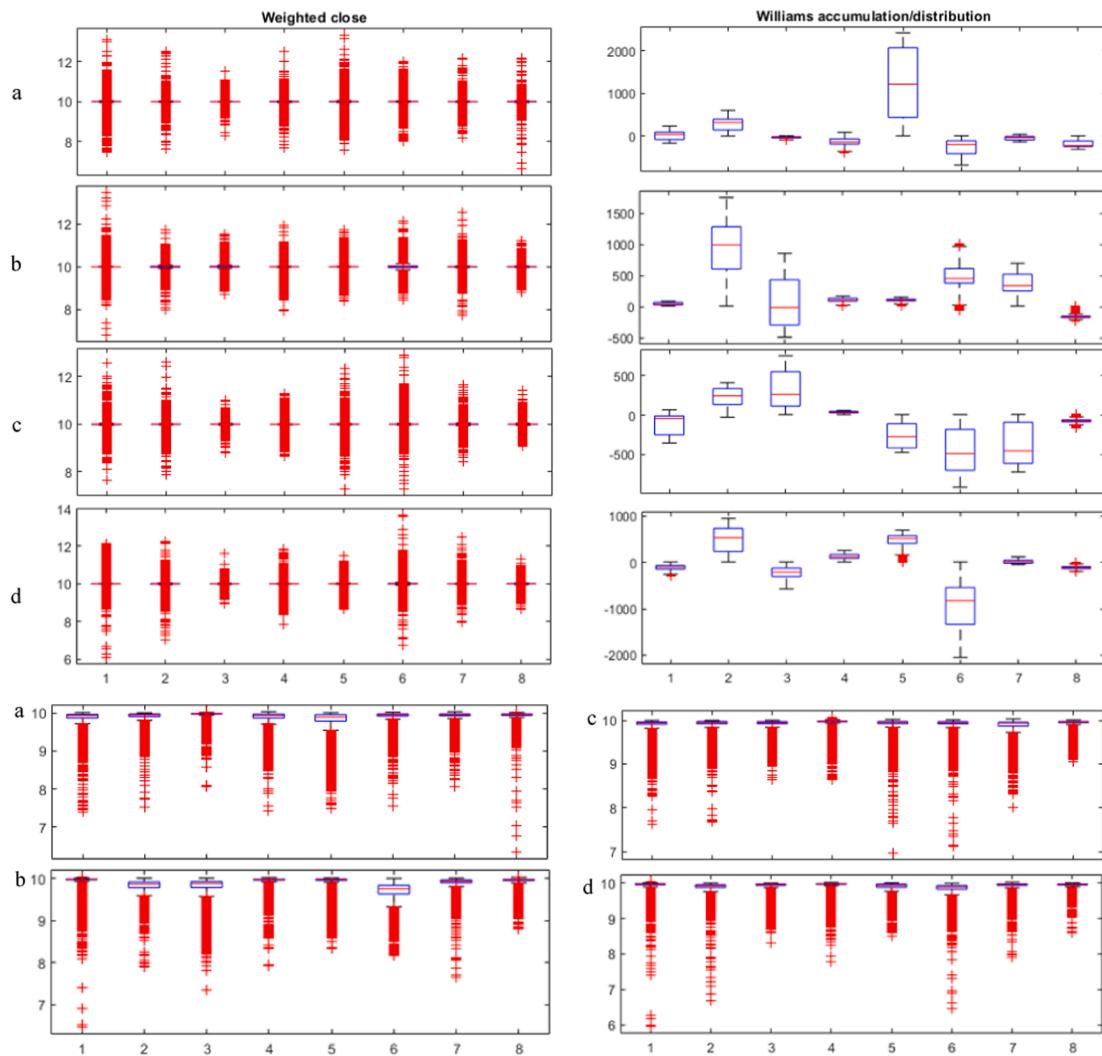


Fig. C1. (continued).

Appendix D.: Dataset details

The details of the datasets are provided below:

- HD-sEMG [37,57]: The HD-sEMG dataset¹ contains recordings from twenty subjects performing 34 hand gestures twice on separate days. 256 electrodes record data at the rate of 2048 Hz in dynamic (1 s duration, from subjects' relaxed state to the required gesture) and maintenance (4 s duration, from subjects' relaxed state to the required gesture followed by maintenance at that gesture) tasks in raw and pre-processed signals. In the pre-processing stage, a 10–500 Hz 8-order Butterworth bandpass filter and a notch filter are applied to the signals.
- EMG [58]: The EMG dataset² contains recordings from thirty subjects at a sample rate of 3000 Hz (here resampled to 1000 Hz) from eight electrodes in seven forearm gestures, including hand open, hand close, wrist flexion, wrist extension, supination, pronation, and rest [45]. Each gesture is repeated four times in random order and paused for three seconds. This dataset demonstrated how to assess sequence-to-sequence classification for identifying various arm movements using EMG signals.³ For BioTF, we generate a sequence of labels for a series of frames by selecting the last label within each related frame to capture the most recent state. Here, we focus on the five more conventional gestures and exclude S and P.
- EEG [59]: The EEG dataset⁴ contains 2519 EEG recordings clinically classified as normal or abnormal, including 6000 sample-long from 21 electrodes. 2284 recordings are used for training and 235 recordings for testing. A majority of the EEG data is sampled at 250 Hz. However, we

¹ <https://physionet.org/content/hd-semg/1.0.0/>.

² Available at <https://ssd.mathworks.com/supportfiles/SPT/data/MyoelectricData.zip>.

³ <https://www.mathworks.com/help/signal/ug/classify-arm-motions-using-emg-signals-and-deep-learning.html>.

⁴ Available at https://www.isip.piconepress.com/projects/tuh_eeg/html/downloads.shtml. Please note that 474 recording are removed from the main dataset (2993 recording) because of data incompatibility with the training workflow.

resample the data to 100 Hz to prevent high computation needs. Also, values out of the $800 \mu\text{V}$ for reducing the effect of artifacts.

- ECG [60,61]: Here, we use the ECG dataset⁵ from the PhysioNet 2017 Challenge. The dataset includes four classes: normal, atrial fibrillation, other rhythms, and noisy recording, sampled at a frequency of 300 Hz. Here, we focus on classifying normal and atrial fibrillation.

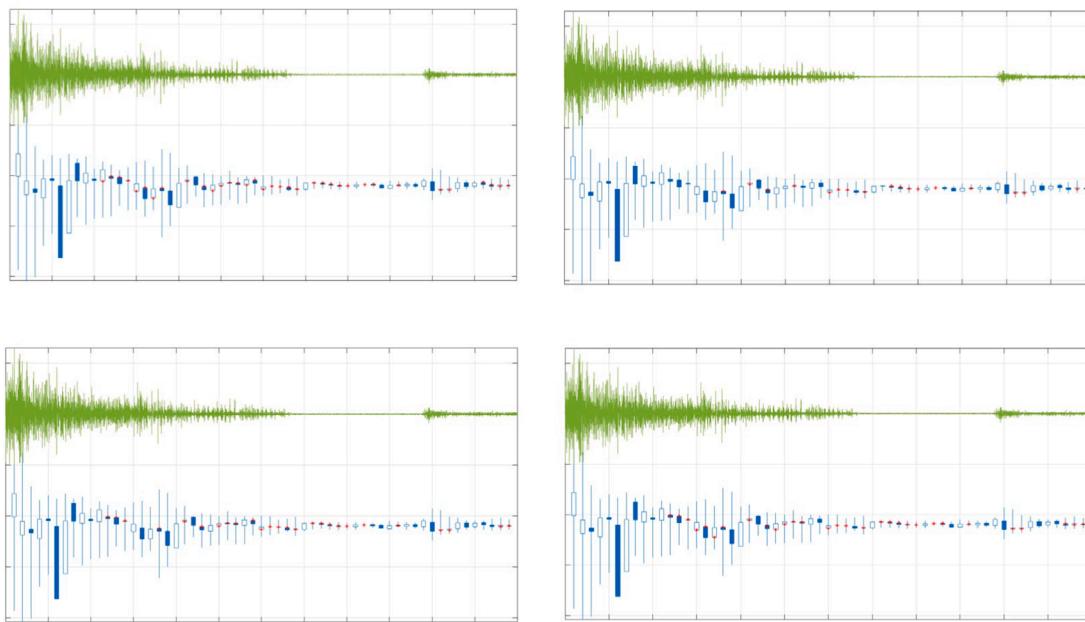


Fig. D1. Financial pattern recognition for EMG signal. The patterns from left to right are Neutral (High Wave, Doji, Rickshaw Man, and Spinning Top Pattern). Red stars mark neutral candlesticks.

Table G1

The network architecture.

| Data set | Type | Activations | Learnable Properties | # of parameters |
|----------|---|---|---|-----------------|
| EMG | Sequence input with 8 dimensions | $8(\text{C}) \times 1(\text{B}) \times 1(\text{T})$ | – | 0 |
| | LSTM with 80 hidden units | $80(\text{C}) \times 1(\text{B}) \times 1(\text{T})$ | Input weights: 320×8 , Recurrent weights: 320×80 , Bias: 320×1 | 28,480 |
| | Fully connected with 5 neurons | $5(\text{C}) \times 1(\text{B}) \times 1(\text{T})$ | Weights: 5×80 , Bias: 5×1 | 405 |
| | Softmax | $5(\text{C}) \times 1(\text{B}) \times 1(\text{T})$ | – | 0 |
| | Cross entropy out | $5(\text{C}) \times 1(\text{B}) \times 1(\text{T})$ | – | 0 |
| EEG | Input image ($21 \times 6000 \times 1$) | $21(\text{S}) \times 6000(\text{S}) \times 1(\text{C}) \times 1(\text{B})$ | – | 0 |
| | $25 1 \times 10$ convolutions with stride [1 1] and padding [0 0 0] | $21(\text{S}) \times 5991(\text{S}) \times 25(\text{C}) \times 1(\text{B})$ | Weights: $1 \times 10 \times 1 \times 25$, bias: $1 \times 1 \times 25$ | 275 |
| | $25 21 \times 1$ convolutions with stride [1 3] and padding [0 0 0] | $1(\text{S}) \times 1997(\text{S}) \times 25(\text{C}) \times 1(\text{B})$ | Weights: $21 \times 1 \times 25 \times 25$, bias: $1 \times 1 \times 25$ | 13,150 |
| | Batch normalization | $1(\text{S}) \times 1997(\text{S}) \times 25(\text{C}) \times 1(\text{B})$ | Offset: $1 \times 1 \times 25$, scale: $1 \times 1 \times 25$ | 50 |
| | Exponential linear unit with alpha 1 | $1(\text{S}) \times 1997(\text{S}) \times 25(\text{C}) \times 1(\text{B})$ | – | 0 |
| | 1×3 max pooling with stride [1 1] and padding [0 0 0] | $1(\text{S}) \times 1995(\text{S}) \times 25(\text{C}) \times 1(\text{B})$ | – | 0 |
| | Identity layer | $1(\text{S}) \times 1995(\text{S}) \times 25(\text{C}) \times 1(\text{B})$ | – | 0 |
| | 50 % dropout | $1(\text{S}) \times 1995(\text{S}) \times 25(\text{C}) \times 1(\text{B})$ | – | 0 |
| | $50 1 \times 10$ convolutions with stride [1 3] and padding [0 0 0] | $1(\text{S}) \times 662(\text{S}) \times 50(\text{C}) \times 1(\text{B})$ | Weights: $1 \times 10 \times 25 \times 50$, bias: $1 \times 1 \times 50$ | 12,550 |
| | Batch normalization | $1(\text{S}) \times 662(\text{S}) \times 50(\text{C}) \times 1(\text{B})$ | Offset: $1 \times 1 \times 50$, scale: $1 \times 1 \times 50$ | 100 |
| | Exponential linear unit with alpha 1 | $1(\text{S}) \times 662(\text{S}) \times 50(\text{C}) \times 1(\text{B})$ | – | 0 |
| | 1×3 maxpooling with stride [1 1] and padding [0 0 0] | $1(\text{S}) \times 660(\text{S}) \times 50(\text{C}) \times 1(\text{B})$ | – | 0 |
| | Identity layer | $1(\text{S}) \times 660(\text{S}) \times 50(\text{C}) \times 1(\text{B})$ | – | 0 |

(continued on next page)

⁵ Available at <https://physionet.org/content/challenge-2017/1.0.0/>.

Table G1 (continued)

| Data set | Type | Activations | Learnable Properties | # of parameters |
|----------|---|-----------------------------------|---|-----------------|
| ECG | 50 % dropout | 1(S) × 660(S) × 50(C) × 1 (B) | — | 0 |
| | 100 1 × 10 convolutions with stride [1 3] and padding [0 0 0 0] | 1(S) × 217(S) × 100(C) × 1 (B) | Weights: 1 × 10 × 50 × 100, bias: 1 × 1 × 100 | 50,100 |
| | Batch normalization | 1(S) × 217(S) × 100(C) × 1 (B) | Offset: 1 × 1 × 100, scale: 1 × 1 × 100 | 200 |
| | Exponential linear unit with alpha 1 | 1(S) × 217(S) × 100(C) × 1 (B) | — | 0 |
| | 1 × 3 maxpooling with stride [1 1] and padding [0 0 0] | 1(S) × 215(S) × 100(C) × 1 (B) | — | 0 |
| | Identity layer | 1(S) × 215(S) × 100(C) × 1 (B) | — | 0 |
| | 50 % dropout | 1(S) × 215(S) × 100(C) × 1 (B) | — | 0 |
| | 200 1 × 10 convolutions with stride [1 3] and padding [0 0 0 0] | 1(S) × 69(S) × 200(C) × 1 (B) | Weights: 1 × 10 × 100 × 200, bias: 1 × 1 × 200 | 200,200 |
| | Batch normalization | 1(S) × 69(S) × 200(C) × 1 (B) | Offset: 1 × 1 × 200, scale: 1 × 1 × 200 | 400 |
| | Exponential linear unit with alpha 1 | 1(S) × 69(S) × 200(C) × 1 (B) | — | 0 |
| | 1 × 3 maxpooling with stride [1 1] and padding [0 0 0] | 1(S) × 67(S) × 200(C) × 1 (B) | — | 0 |
| | Identity layer | 1(S) × 67(S) × 200(C) × 1 (B) | — | 0 |
| | 2 1 × 67 convolutions with stride [1 1] and padding [0 0 0] | 1(S) × 1(S) × 2(C) × 1(B) | Weights: 1 × 67 × 200 × 2, bias: 1 × 1 × 2 | 26,802 |
| | Softmax | 1(S) × 1(S) × 2(C) × 1(B) | — | 0 |
| | Cross entropy output | 1(S) × 1(S) × 2(C) × 1(B) | — | 0 |
| | Sequence input with one dimension | 1(C) × 1(B) × 1(T) | — | 0 |
| | BiLSTM with 50 hidden units | 100(C) × 1(B) | Input weights: 400 × 1, Recurrent weights: 400 × 50, Bias: 400 × 1 | 20,800 |
| | Fully connected with four neurons | 2(C) × 1(B) | Weight: 2 × 100, Bias: 2 × 1 | 202 |
| | Softmax | 2(C) × 1(B) | — | 0 |
| | Cross entropy output | 2(C) × 1(B) | — | 0 |

Appendix G.: Network architecture

Table G.1 shows the deep learning architecture of the traditional data-point-based classification used for these datasets, and **Table G.2** shows the experiment hyperparameters. Only the input layers are changed to fit data for time frame-based processing. All networks are trained using MATLAB (2022b) built-in ADAM optimizer. We use built-in hyperparameter optimization for the LDA structure, minimizing five-fold cross-validation loss.

Appendix H.: Feature sets

Table H1
Feature sets details.

| Feature set name | Features |
|------------------|----------------------|
| D1 | RMS, WL, ZC, SSC |
| D2 | RMS |
| D3 | RMS, WL |
| D4 | RMS, WL, ZC |
| D5 | WL |
| D6 | ZC |
| D7 | SSC |
| D8 | ZC, SSC |
| F1 | FBL based on $w = 5$ |

Table G2

Hyperparameters of experiments.

| Dataset | Hyperparameters | | | | |
|---------|-----------------------|------------|-------|--------------------|------|
| | Initial learning rate | Batch size | Epoch | Gradient threshold | Test |
| EMG | 0.001 | 8 | 100 | 1 | 20 % |
| EEG | 0.001 | 8 | 35 | — | 10 % |
| ECG | 0.001 | 20 | 500 | 1 | 10 % |

The data point-based feature set includes root mean squared (RMS), wavelength (WL), ZC, and SSC [37]. The frame-based feature set includes frame body length (FBL). The frame body length means the summation of the absolute difference of starting and ending values of frames. More details are provided in Table H.1.

Appendix I . Explanation's parameters

LIME and input node activation explanations of the networks are provided using *imageLIME* and *activations* built-in MATLAB functions, respectively. Table I.1 shows the parameters of the LIME explanation. It is recommended that readers read the MATLAB documentation for more details.

Table I.1
Parameters of explanation methods.

| LIME Property | Value |
|------------------------------------|--------|
| Synthesis data | Global |
| Number of synthesis data | 5000 |
| Number of query point neighbors | 1500 |
| Simple model type | Linear |
| Tolerance of linear coefficients | 1e-4 |
| Kernel width | 0.75 |
| Minkowski distance matrix exponent | 2 |

References

- [1] E. Tjoa, C. Guan, A survey on explainable artificial intelligence (xai): toward medical xai, *IEEE Trans. Neural Networks Learn. Syst.* 32 (11) (2020) 4793–4813.
- [2] H. W. Loh, C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, and U. R. Acharya, “Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022),” *Computer Methods and Programs in Biomedicine*, p. 107161, 2022.
- [3] S. Cho, G. Lee, W. Chang, J. Choi, Interpretation of deep temporal representations by selective visualization of internally activated nodes, “arXiv preprint arXiv: 2004.12538, 2020.
- [4] G. Montavon, S. Lapuschkin, A. Binder, W. Samek, K.-R. Müller, Explaining nonlinear classification decisions with deep Taylor decomposition, *Pattern Recogn.* 65 (2017) 211–222.
- [5] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, Learning deep features for discriminative localization, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2921–2929.
- [6] T. Rojat, R. Puget, D. Filliat, J. Del Ser, R. Gelin, and N. Díaz-Rodríguez, “Explainable artificial intelligence (xai) on time series data: A survey,” *arXiv preprint arXiv:2104.00950*, 2021.
- [7] Z. Wang, W. Yan, and T. Oates, “Time series classification from scratch with deep neural networks: A strong baseline,” in *2017 International joint conference on neural networks (IJCNN)*, 2017: IEEE, pp. 1578–1585.
- [8] S.A. Siddiqui, D. Mercier, M. Munir, A. Dengel, S. Ahmed, Tsviz: Demystification of deep learning models for time-series analysis, *IEEE Access* 7 (2019) 67027–67040.
- [9] S. Tonekaboni, S. Joshi, D. Duvenaud, and A. Goldenberg, “Explaining time series by counterfactuals,” 2019. [Online]. Available: <https://openreview.net/forum?id=HygDF1rYDB>.
- [10] D. Bau, B. Zhou, A. Khosla, A. Oliva, and A. Torralba, “Network dissection: Quantifying interpretability of deep visual representations,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 6541–6549.
- [11] N. Strothoff, C. Strothoff, Detecting and interpreting myocardial infarction using fully convolutional neural networks, *Physiol. Meas.* 40 (1) (2019), 015001.
- [12] A. Theissler, F. Spinnato, U. Schlegel, R. Guidotti, Explainable AI for time series classification: a review, taxonomy and research directions, *IEEE Access* 10 (2022) 100700–100724.
- [13] H. Lee, D. Kim, Y.-L. Park, Explainable deep learning model for EMG-based finger angle estimation using attention, *IEEE Trans. Neural Syst. Rehabil. Eng.* 30 (2022) 1877–1886.
- [14] T.-Y. Hsieh, S. Wang, Y. Sun, V. Honavar, Explainable multivariate time series classification: a deep neural network which learns to attend to important variables as well as time intervals, in: Proceedings of the 14th ACM international conference on web search and data mining, 2021, pp. 607–615.
- [15] D. Raab, A. Theissler, M. Spiliopoulou, XAI4EEG: spectral and spatio-temporal explanation of deep learning-based seizure detection in EEG time series, *Neural Comput. & Appl.* (2022) 1–18.
- [16] L. Lin, et al., An EEG-based cross-subject interpretable CNN for game player expertise level classification, *Expert Syst. Appl.* 237 (2024), 121658.
- [17] N. Alamatsaz, L. Tabatabaei, M. Yazdchi, H. Payan, N. Alamatsaz, F. Nasimi, A lightweight hybrid CNN-LSTM explainable model for ECG-based arrhythmia detection, *Biomed. Signal Process. Control* 90 (2024), 105884.
- [18] M.J. Nkengue, X. Zeng, L. Koehl, X. Tao, X-RCRNet: An explainable deep-learning network for COVID-19 detection using ECG beat signals, *Biomed. Signal Process. Control* 87 (2024), 105424.
- [19] K. Qin, W. Huang, T. Zhang, H. Zhang, X. Cheng, A lightweight SelfONN model for general ECG classification with pretraining, *Biomed. Signal Process. Control* 89 (2024), 105780.
- [20] H. Abbass, K. Crockett, J. Garibaldi, A. Gegov, U. Kaymak, J.M.C. Sousa, From explainable artificial intelligence (xAI) to understandable artificial intelligence (uAI), *IEEE Trans. Artif. Intell.* 5 (09) (2024) 4310–4314.
- [21] A.A. Torres-García, O. Mendoza-Montoya, M. Molinas, J. M. Antelis, L. A. Moctezuma, T. Hernández-Del-Toro, Pre-processing and feature extraction, in: *Biosignal processing and classification using computational learning and intelligence*, Elsevier, 2022, pp. 59–91.
- [22] M.A. Oskooi, H. Hu, Myoelectric control systems—A survey, *Biomed. Signal Process. Control* 2 (4) (2007) 275–294.
- [23] M. Zecca, S. Micera, M. C. Carrozza, P. Dario, Control of multifunctional prosthetic hands by processing the electromyographic signal, “Critical Reviews™ in Biomedical Engineering, vol. 30, no. 4–6, 2002.
- [24] Z. Liu, J. Sun, Y. Zhang, P. Rolfe, Sleep staging from the EEG signal using multi-domain feature extraction, *Biomed. Signal Process. Control* 30 (2016) 86–97.
- [25] H.J. Escalante, E.F. Morales, “Dimensionality reduction,” in *Biosignal processing and classification using computational learning and intelligence*, Elsevier (2022) 93–108.
- [26] X. Jiang, K. Nazarpour, C. Dai, Explainable and robust deep forests for EMG-force modeling, *IEEE J. Biomed. Health Informatics*, 2023.
- [27] A. Phinyomark, P. Phukpattaranont, C. Limsakul, Feature reduction and selection for EMG signal classification, *Expert Syst. Appl.* 39 (8) (2012) 7420–7431.
- [28] C. Spiewak, M. Islam, A. Zaman, M.H. Rahman, A comprehensive study on EMG feature extraction and classifiers, *Open Access Journal of Biomedical Engineering and Biosciences* 1 (1) (2018) 1–10.
- [29] S. Nison, The Candlestick Course-Steve Nison, ed: John Wiley & Sons, Inc, Hoboken, New Jersey, 2011.
- [30] J.C. Brammer, Biopeaks: A graphical user interface for feature extraction from heart-and breathing biosignals, *Journal of Open Source Software* 5 (54) (2020) 2621.
- [31] L.A. Kallenberg, E. Schulte, C. Disselhorst-Klug, H.J. Hermens, Myoelectric manifestations of fatigue at low contraction levels in subjects with and without chronic pain, *J. Electromogr. Kinesiol.* 17 (3) (2007) 264–274.
- [32] I.J.R. Martínez, A. Mannini, F. Clemente, C. Cipriani, Online grasp force estimation from the transient EMG, *IEEE Trans. Neural Syst. Rehabil. Eng.* 28 (10) (2020) 2333–2341.
- [33] D. D'Accolti, K. Dejanovic, L. Cappello, E. Mastinu, M. Ortiz-Catalan, C. Cipriani, Decoding of multiple wrist and hand movements using a transient EMG classifier, *IEEE Trans. Neural Syst. Rehabil. Eng.* (2022).
- [34] G. Kanitz, C. Cipriani, B.B. Edin, Classification of transient myoelectric signals for the control of multi-grasp hand prostheses, *IEEE Trans. Neural Syst. Rehabil. Eng.* 26 (9) (2018) 1756–1764.
- [35] M.T. Ribeiro, S. Singh, C. Guestrin, Why should I trust you?“ Explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1135–1144.
- [36] C.-K. Chen, C.-L. Lin, S.-L. Lin, Y.-M. Chiu, C.-T. Chiang, A chaotic theoretical approach to ECG-based identity recognition [application notes], *IEEE Comput. Intell. Mag.* 9 (1) (2014) 53–63.
- [37] X. Jiang, et al., Open access dataset, toolbox and benchmark processing results of high-density surface electromyogram recordings, *IEEE Trans. Neural Syst. Rehabil. Eng.* 29 (2021) 1035–1046.
- [38] A. Asghar, S. Jawaid Khan, F. Azim, C. S. Shakeel, A. Hussain, I.K. Niazi, Review on electromyography based intention for upper limb control using pattern recognition for human-machine interaction, Proceedings of the Institution of Mechanical

- Engineers, Part H: Journal of Engineering in Medicine, vol. 236, no. 5, pp. 628-645, 2022.
- [39] T. Tuncer, S. Dogan, U.R. Acharya, Automated EEG signal classification using chaotic local binary pattern, *Expert Syst. Appl.* 182 (2021), 115175.
- [40] E. Khalili, B.M. Asl, Automatic sleep stage classification using temporal convolutional neural network and new data augmentation technique from raw single-channel EEG, *Comput. Methods Programs Biomed.* 204 (2021), 106063.
- [41] S.W. Chen, S.L. Wang, X.Z. Qi, S.M. Samuri, C. Yang, Review of ECG detection and classification based on deep learning: Coherent taxonomy, motivation, open challenges and recommendations, *Biomed. Signal Process. Control* 74 (2022), 103493.
- [42] J. Yoo, T.J. Jun, Y.-H. Kim, xECGNet: Fine-tuning attention map within convolutional neural network to improve detection and explainability of concurrent cardiac arrhythmias, *Comput. Methods Programs Biomed.* 208 (2021), 106281.
- [43] M. Kiani, J. Andreu-Perez, H. Hagras, S. Rigato, M.L. Filippetti, Towards understanding human functional brain development with explainable artificial intelligence: challenges and perspectives, *IEEE Comput. Intell. Mag.* 17 (1) (2022) 16–33.
- [44] Z. Li, M. Hayashibe, C. Fattal, D. Guiraud, Muscle fatigue tracking with evoked EMG via recurrent neural network: Toward personalized neuroprosthetics, *IEEE Comput. Intell. Mag.* 9 (2) (2014) 38–46.
- [45] A.R. Goge, A. Chan, Investigating classification parameters for continuous myoelectrically controlled prostheses, CMBES Proceedings, vol. 28, 2005.
- [46] S. López, I. Obeid, J. Picone, Automated interpretation of abnormal adult electroencephalograms, Temple University, Libraries, 2017.
- [47] A. Anand, T. Kadian, M.K. Shetty, A. Gupta, Explainable AI decision model for ECG data of cardiac disorders, *Biomed. Signal Process. Control* 75 (2022), 103584.
- [48] G.R. Patra, S.K. Mohapatra, M.N. Mohanty, Applications of deep learning algorithms in biomedical signal processing-pros and cons, *Int. J. Biomet.* 14 (1) (2022) 98–124.
- [49] A. Vaswani, et al., Attention is all you need, *Adv. Neural Inf. Proces. Syst.* 30 (2017).
- [50] R.N. Khushaba, A.H. Al-Timemy, O.W. Samuel, E.J. Scheme, Myoelectric control with fixed convolution-based time-domain feature extraction: exploring the spatio-temporal interaction, *IEEE Trans. Hum.-Mach. Syst.* 52 (6) (2022) 1247–1257.
- [51] A. Supratak, C. Wu, H. Dong, K. Sun, Y. Guo, Survey on feature extraction and applications of biosignals, in: Machine learning for health informatics: Springer, 2016, pp. 161–182.
- [52] H.P. Martinez, Y. Bengio, G.N. Yannakakis, Learning deep physiological models of affect, *IEEE Comput. Intell. Mag.* 8 (2) (2013) 20–33.
- [53] H. Shi, X. Jiang, C. Dai, W. Chen, EMG-based multi-user hand gesture classification via unsupervised transfer learning using unknown calibration gestures, *IEEE Trans. Neural Syst. Rehabil. Eng.* (2024).
- [54] G. Hajian, et al., Generalizing upper limb force modeling with transfer learning: a multimodal approach using EMG and IMU for new users and conditions, *IEEE Trans. Neural Syst. Rehabil. Eng.* (2024).
- [55] A. Moradi, et al., Clinical implementation of a bionic hand controlled with kinetico-myographic signals, *Sci. Rep.* 12 (1) (2022/08/31 2022), 14805, <https://doi.org/10.1038/s41598-022-19128-1>.
- [56] A. Abasian, H. Rafiei, M.-R. Akbarzadeh-T., A. Akbarzadeh, A. Moradi, A.-M. Naddaf-Sh, Stochastic Sequential Sensory Selection for Gesture Recognition in KineticoMyoGraphy Guided Bionic Hands, *IEEE Transactions on Medical Robotics and Bionics* (2024), <https://doi.org/10.1109/TMRB.2024.3503993>.
- [57] X. Jiang, C. Dai, X. Liu, J. Fan, Open access dataset and toolbox of high-density surface electromyogram recordings, PhysioNet, 2021.
- [58] A.D. Chan, G.C. Green, Myoelectric control development toolbox, CMBES Proceedings, vol. 30, 2007.
- [59] I. Obeid, J. Picone, The temple university hospital EEG data corpus, *Front. Neurosci.* 10 (2016), 195498.
- [60] G.D. Clifford et al., AF classification from a short single lead ECG recording: the physionet/computing in cardiology challenge 2017, in: 2017 Computing in Cardiology (CinC), 2017: IEEE, pp. 1–4.
- [61] A.L. Goldberger, et al., PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, *Circulation* 101 (23) (2000) e215–e220.