

Article

Synthesis of sEMG Signals for Hand Gestures Using a 1DDCGAN

Mohamed Amin Gouda , Wang Hong *, Daqi Jiang, Naishi Feng, Bin Zhou and Ziyang Li 

Department of Mechanical Engineering and Automation, Northeastern University, Shenyang 110819, China; mohamedamin@stumail.neu.edu.cn (M.A.G.); 1910103@stu.neu.edu.cn (D.J.); fengnaishi98@163.com (N.F.); 1810097@stu.neu.edu.cn (B.Z.); 2110108@stu.neu.edu.cn (Z.L.)

* Correspondence: hongwang@mail.neu.edu.cn

Abstract: The emergence of modern prosthetics controlled by bio-signals has been facilitated by AI and microchip technology innovations. AI algorithms are trained using sEMG produced by muscles during contractions. The data acquisition procedure may result in discomfort and fatigue, particularly for amputees. Furthermore, prosthetic companies restrict sEMG signal exchange, limiting data-driven research and reproducibility. GANs present a viable solution to the aforementioned concerns. GANs can generate high-quality sEMG, which can be utilised for data augmentation, decrease the training time required by prosthetic users, enhance classification accuracy and ensure research reproducibility. This research proposes the utilisation of a one-dimensional deep convolutional GAN (1DDCGAN) to generate the sEMG of hand gestures. This approach involves the incorporation of dynamic time wrapping, fast Fourier transform and wavelets as discriminator inputs. Two datasets were utilised to validate the methodology, where five windows and increments were utilised to extract features to evaluate the synthesised sEMG quality. In addition to the traditional classification and augmentation metrics, two novel metrics—the Mantel test and the classifier two-sample test—were used for evaluation. The 1DDCGAN preserved the inter-feature correlations and generated high-quality signals, which resembled the original data. Additionally, the classification accuracy improved by an average of 1.21–5%.

Keywords: DCGAN; EMG; sEMG; AI; bio-signals; data augmentation; classification



Citation: Gouda, M.A.; Hong, W.; Jiang, D.; Feng, N.; Zhou, B.; Li, Z. Synthesis of sEMG Signals for Hand Gestures Using a 1DDCGAN. *Bioengineering* **2023**, *10*, 1353. <https://doi.org/10.3390/bioengineering10121353>

Academic Editors: Jiejunyi Liang, Ting Zhang and Jianjun Meng

Received: 7 October 2023

Revised: 18 November 2023

Accepted: 20 November 2023

Published: 25 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Muscle contractions create electromyography (EMG) signals, which can be used to measure muscle activity and to diagnose muscle disorders. These EMG signals can be captured using surface, needle and intramuscular electrodes [1–3]. However, the prevalent approach in research and prosthetic applications involves the utilisation of surface electromyography (sEMG) due to its ability to capture global muscle information, in contrast to needles, which only gather data from a single muscle. Moreover, the utilisation of needles and intramuscular electrodes presents potential drawbacks, such as pain and inherent risks.

Various privacy laws may impose restrictions on researchers who require access to sEMG signals from specialised companies. Researchers may find it challenging to carry out repeatable research and product development as a result, which limits replicable research and product development. Researchers created multiple datasets, including those mentioned in Refs [4,5] and others, to provide independent data sources. Nevertheless, these datasets might have missing or damaged data [6] or lack the gesture repetitions needed to build a robust classification model, resulting in overfitting, according to Kaczmarek et al. [7]. Some researchers represented the sEMG signal in two dimensions and employed translation and rotation to augment the data, but these methods are not appropriate for one-dimensional time series data [8].

In order to overcome these issues, generative adversarial networks (GAN) can be utilised to generate new data instances. There are various methods to generate data.

For example, general-purpose parametric models can also be used to create data in the easiest way [9]; however, they are the least accurate, and these signals can be immediately identified as fabricated data. Expert-driven parametric models are more accurate but require more experience. Anonymised bio-signals can be used to deliver data in a private manner. For both the control and diseased groups, it is imperative to adhere to experimental best practices when recording, processing and storing signals. Data re-identification is a possibility if appropriate collecting procedures are not followed. Deep-learning models such as GAN have been used to synthesise bio-signals by learning the statistical patterns of the original data and producing synthetic data. GANs demand deep-learning expertise but little domain knowledge.

GANs can let organisations share synthesised signals with researchers without breaking patients' data privacy laws, boosting research repeatability. All approaches must protect patient privacy and follow privacy standards. The synthesised data have the potential to be utilised for the purpose of training AI-controlled prostheses, thereby enabling the achievement of precise control. Typically, in the context of an amputee using an AI-controlled prosthetic device, sEMG electrodes are attached to the individual's muscles. Subsequently, the amputee proceeds to execute various hand gestures, with this iterative process being repeated multiple times until the AI model achieves convergence. The act of iteratively performing the gesture on multiple occasions is exhausting and introduces complexity to the overall procedure. GANs have the potential to expedite the convergence of the classification model by generating synthetic data. Furthermore, GANs may potentially avoid the necessity for amputees to visit a laboratory, allowing them to perform the process in the comfort of their own homes. Data synthesis protects patient privacy and enhances research.

Related Work

GANs were initially proposed by Ian Goodfellow; it is a type of machine-learning (ML) model, which can be used to generate realistic data. GANs work by pitting two neural networks against each other: a generator, which creates new data, and a discriminator, which tries to distinguish between real and fake data [10].

Research on GANs has grown significantly, with early work focusing on generating images [10,11]. Later, researchers began to focus on generating text [12] and speech [13]. More recently, researchers have begun to consider using GANs to generate data, which are not easily accessible, such as medical images or financial data. The generation of medical data has become increasingly important with the advancement in pattern recognition techniques and the rise in privacy concerns. Generating medical data focuses on finding new deep-learning models, which can generate quality data and ways to evaluate them. Esteban et al. proposed RGAN and RCGAN time series data models [14]. Visual and quantitative methods, such as sample likelihood and maximum mean discrepancy (MMD), were used to examine the data. Train on synthetic, test on real (TSTR) and train on real, test on synthetic (TRTS) are novel evaluation methods, which use the generated data to train a classifier and the actual data to evaluate it and vice versa. The authors designed three tests—one qualitative and two statistical—to mitigate overfitting due to privacy concerns regarding sensitive material.

Hazra et al. [15] developed a one-dimensional GAN called SynSigGan, which works well on biological signals. Their model refines signals using discrete wavelet transform, thresholding and inverse discrete wavelet transform and then evaluates the generated data using the Pearson correlation coefficient, root mean square (RMS) error, per cent RMS difference, mean absolute error and Fréchet distance (FD) for statistical analysis. Researchers suggested a unique method employing bidirectional RNN and statistical stages to generate synthetic biological signals for patients or events. The method generates five biological signals and outperforms other generative models in evaluation metrics.

Beaulieu-Jones et al. [16] noted that sharing individual-level clinical study data is hindered by challenging requirements. The sharing of data among academics is often hindered by the need for formal collaborations and elaborate data usage agreements. Deep neural

networks were combined to create a systolic blood pressure trial (SPRINT). Differential privacy, which reduces the likelihood of identifying clinical research participants, was used to train neural networks. Hypothesis-generating analyses simulated the original trial data using machine-learning predictors from the synthetic population. The discovery suggests that synthetic data enable secondary studies and replicable exploration of clinical datasets while protecting participant confidentiality.

Anicet Zanini et al. [17], in an effort to solve the lack of available Parkinson's data, proposed approaches based on DCGAN and a combination of DCGAN and style, introducing two new data augmentation methods using DCGANs and style transfer for Parkinson's disease sEMG signals. The proposed models can simulate individual patients' tremor patterns and extend them to different movement protocols. This allows for the extension of patient datasets and the generation of tremor simulations for validating treatment approaches in different movement scenarios. The experimental results demonstrate the models' ability to adapt to various frequencies and amplitudes of tremors, making them a promising tool for Parkinson's disease research. Moreover, fast Fourier transform (FFT), mean square error, dynamic time wrapping (DTW) and sEMG envelope cross-correlation are the metrics used in their research to evaluate the generated signals. The mean square error (MSE) of FFT magnitudes is used to measure the similarity between time series signals. DTW is used in time series analysis to measure the similarity between sequences by comparing local cost functions. Cross-correlation measures the similarity between two series by considering their displacement.

Campbell et al. [18] developed a framework for brief training programmes using subject-specific synthetic sEMG data using SinGAN to generate synthetic data. When combined with restricted training, artificial data can improve classification accuracy. The study created 1000 fake sEMG segments comprising six actions. Qualitative, quantitative and classification methods examined the results. Artificially created data improved classification accuracy by 5.4% in cross-validation testing. This method could improve myoelectric control training in data-poor situations.

Zhang et al. [8] used one-dimensional energy-based generative adversarial networks (EBGAN) to generate sEMG features to improve classification precision. The discriminant network uses the energy paradigm instead of binary assessment, and the fully linked layers capture the distribution of genuine sEMG data to create comparable data. This GAN design achieves a lower MMD than others. This data augmentation method increased traditional classification model precision by 1.21–5%.

In brief, numerous studies have been conducted in the field of GANs, with a particular emphasis on their application in image generation. Subsequently, GANs have been employed for the analysis of various types of data, including bio-signals. Many studies approach time series bio-signals as two-dimensional images, which may not be suitable for this particular type of data. Furthermore, it is crucial to contemplate the establishment of appropriate metrics for the said signals.

To this end, this research employs a 1DDCGAN with DTW, FFT and wavelets as inputs to the discriminator to generate the synthetic sEMG of hand gestures. This method generates synthetic data to expand the datasets. Two different sEMG datasets were used in this research to validate this approach. In order to evaluate the quality of data generated by the model, several evaluation techniques were employed, including the classifier two-sample (C2ST) test, the Mantel test and conventional classification metrics, such as classification accuracy and augmentation test. The proposed model improved the classification accuracy by 1.21–5% in the augmentation test. The two novel metrics showed that the model produced high-quality data, which resembled the original data and kept feature interrelationships. To the author's knowledge, the usage of 1DDCGAN, the C2ST and the Mantel test is new in the context of EMG hand gestures.

This research paper is divided as follows:

- Section 1: Introduction, which provides the background on sEMG signals and the challenges with acquiring data, as well as an overview of using GANs to generate synthetic medical data.
- Section 2: Materials and Methods, which details the data acquisition, signal processing, proposed 1DDCGAN architecture and evaluation methods, including the Mantel test, classification, augmentation test and classifier two-sample test.
- Section 3: Results, which presents the results of the Mantel test, classification, augmentation test and classifier two-sample test in assessing the quality of the synthesised sEMG signals.
- Section 4: Discussion, which analyses and interprets the results, comparing them with previous literature.
- Section 5: Conclusion, which summarises the efficacy of using 1DDCGAN to generate synthetic sEMG signals based on the evaluation metrics.

2. Materials and Methods

The objective of this study is to create a GAN, which can produce high-quality signals capable of reproducing the characteristics of the original hand gesture signal, as well as augmenting the datasets, which lack repetitions or missing data points to improve the classification. Figure 1 illustrates the proposed model's flow chart. This section will provide a comprehensive overview of the techniques employed in conjunction with our proposed strategy.

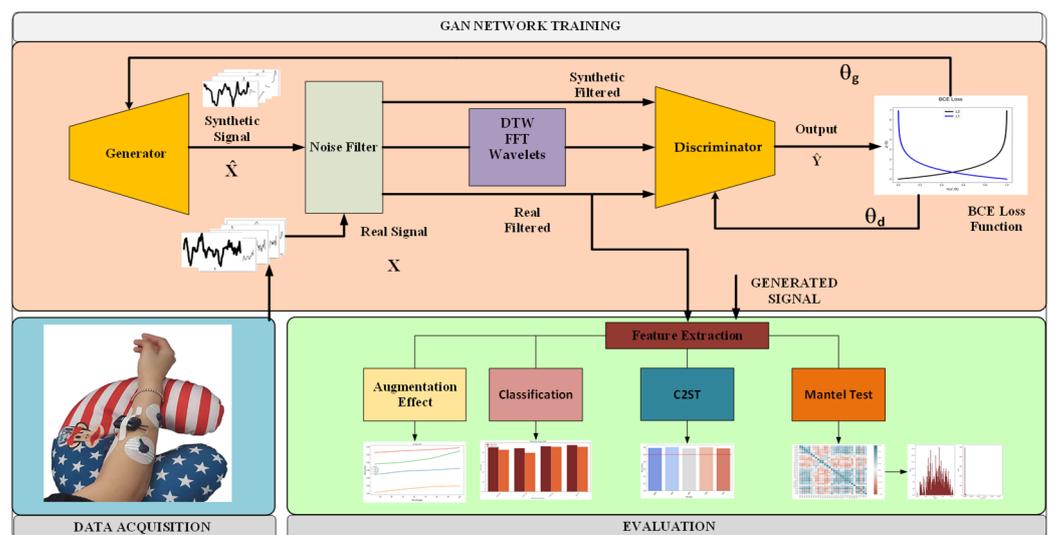


Figure 1. Flow chart of the study. The data acquisition stage (lower left) is followed by a GAN training stage in the top section. The evaluation stage (lower right) evaluates the data quality using different techniques.

Figure 1 depicts the flow chart representing the methodology employed in this study. The sEMG signals obtained from the respective hand gestures and the synthetic signals generated by the generator are initially subjected to noise filtration. Subsequently, the DTW, wavelets and FFT are extracted from both signals. These extracted features, along with the original signals, are subsequently input into the discriminator. Finally, after the training of the 1DDCGAN, the generated signals and the filtered original signals are evaluated using the augmentation classification method to measure the improvement the generated signals provide, classification to examine the accuracy of the generated and original dataset independently, C2ST to examine the ability of pattern recognition to distinguish between the two datasets and the Mantel test to examine whether the relationship between features is preserved. The data acquisition setup and the devices used are detailed in Section 2.1. Signal filtering and the extracted features are explained in Section 2.2. The proposed

architecture and the GAN network theory are detailed in Section 2.3. Finally, the evaluation methods employed are explained in Section 2.4.

2.1. Data Acquisition

In this study, two distinct datasets were used to assess our approach's efficacy in synthesising various sEMG classes. The first dataset was private, and the second was public [19]. Both datasets were collected with two electrodes, whereas the current trend uses an array [4,5,7,20,21]. The usage of fewer electrodes is better in some applications because it can help reduce hardware and computation costs, which can be an essential consideration in practical settings.

For the private dataset, five non-disabled participants were instructed to perform six different gestures. They were two male and three female volunteers between the ages of 26 and 29 and were right-handed. The participants sat with their arms extended on a disc during the recording. Participants were briefed before the experiment. A video showed participants when to hold and release movements during the session. In order to prevent overfitting and reduce variation, the participants were asked to perform different gestures in a sequence rather than repeating the same gesture multiple times consecutively [7]. Each session included six motions, held for three seconds, followed by three seconds of rest; this procedure was repeated 15 times over three days to avoid muscle fatigue. Figure 2 shows the six gestures, the electrode placement and the timeline of the experiment. The apparatus used to collect the private dataset was Myon Atkos-mini. It contained two wireless sEMG sensors with a 2000 Hz sampling frequency. The electrodes were placed on the flexor and extensor digitorum muscles. The experiment was conducted in the Mechatronics laboratory at Northeastern University in Liaoning, China. The procedures of the experiment were in accordance with the university's ethical and other relevant committees, and the participants also provided written consent.

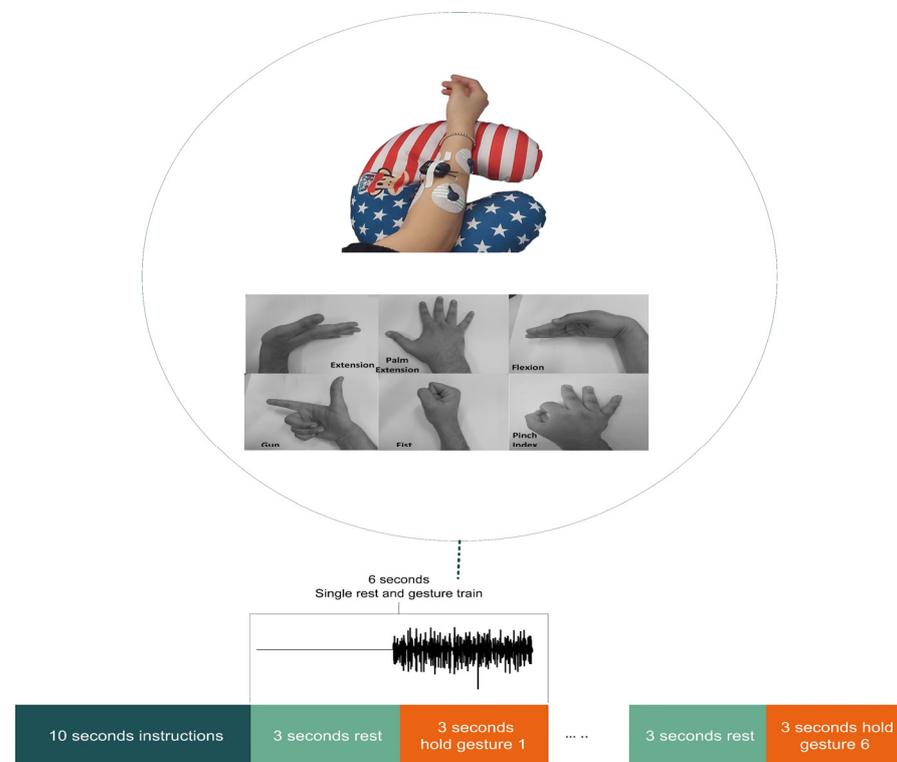


Figure 2. Electrodes' placement during an acquisition session and the six gestures for the private dataset. The timeline of a typical session is 3 s for rest and 3 s for gestures, each repeated 15 times.

The publicly available dataset used in this research was the dataset from Ref [19]. Six men and two women between 20 and 35 years old participated. The subjects were asked to perform ten different gestures, each held for five seconds, then repeated six times. The sEMG data were collected using two channels of the Delsys DE 2 series with a frequency of 4000 Hz. The sensors were placed on the biceps brachii and triceps brachii muscles.

2.2. Signal Processing

Signal processing techniques are applied to EMG data for multiple purposes, including noise reduction, artefact removal and feature extraction.

2.2.1. Filtering

Before the signal was used in the GAN network, it was filtered, so that the generated signal did not have artefacts and could be easily processed by researchers. A bandpass Butterworth filter of the fourth order with pass range set to 10–500 Hz, as it is known to encompass the frequencies associated with muscle activity [21]. Additionally, a notch filter was tuned to 50 Hz, which is the frequency at which power line interference occurs.

2.2.2. Feature Extraction

After the GAN signal generation and before the evaluation stage, features were extracted using the windowing technique, which divides the signal into smaller, overlapping segments called windows. Windowing is particularly useful when the signal being analysed is non-stationary. Sliding windows of various lengths, notably 200, 400, 600, 800 and 1000 data points, were used in the analysis. The amount of overlap was set at 25%, 50%, 75% and 100% of the window length. Afterwards, features were extracted from these different windows with different overlaps.

The extracted features consist of a collection of features in both the time domain and time-frequency domain [22–26] (Equations (1)–(7)). The time domain features used consist of the mean absolute value (MAV_w), slope sign change (SSC), wavelength (WL), mean absolute value slope (MAVS), histogram (HIST), zero crossing (ZC) and RMS. The marginal discrete wavelet transform (mDWT) was selected as the time-frequency domain feature. SSC is a time domain feature, which represents the number of times the sEMG signal slope changes signs in a specific time window. ZC is a time domain feature, which refers to the number of times the sEMG signal crosses the zero-axis in a time interval, which reflects the frequency of signal values changing signs inside a window. MAV_w is a time domain feature, which is a statistical measure of the absolute magnitude of the sEMG signal.

$$MAV_w(x) = \frac{1}{T} \sum |x_t| \quad (1)$$

WL is a time domain feature, which is a measure of the total length of the sEMG signal over a given window of time and is indicative of the complexity of the signal.

$$WL_w(x) = \sum_{t=2}^T |x_t - x_{t-1}| \quad (2)$$

MAVS is a time domain feature, which refers to the rate of change in the MAV between two adjacent windows.

$$MAVS_w(x) = MAV_{w+1}(x) - MAV_w(x) \quad (3)$$

The integrated absolute value (IAV) is a time domain feature, which measures the area under the absolute value of the sEMG signal curve within a given window of time.

$$IAV_w(x) = \sum_{t=1}^T |x_t| \quad (4)$$

HIST is a time domain feature, which measures the sEMG signal amplitude distribution throughout time. The number of sEMG signal values in each bin is counted to determine HIST.

$$HIST(x) = hist(x_{1:T}, B) \tag{5}$$

RMS is a measure of the amplitude of the sEMG signal over a certain time.

$$RMS_w(x) = \sqrt{\frac{1}{T} \sum_{t=1}^T x_t^2} \tag{6}$$

mDWT is a time-frequency domain feature, which measures the signal’s frequency content at different scales and is obtained through a wavelet transform. The wavelet transform was created with three levels, and Daubechies 7 wavelet, l , is the deepest level of the decomposition.

$$mDWT_l(x) = \sum_{\tau=0}^{T/2^l-1} \left| \sum_{t=1}^T x_t \psi_{l,\tau}(t) \right| \tag{7}$$

$$\psi_{l,\tau}(t) = 2^{-\frac{m}{2}} \psi(2^{-l}t - \tau)$$

2.3. Generative Adversarial Network

GANs are effective machine-learning generative models, which are able to generate realistic objects, which are hard to distinguish from realistic ones. A GAN is composed of two distinct neural networks: the generator and the discriminator. The network’s task is to learn the underlying distribution of a dataset and use that knowledge to generate new data similar to the original data.

For a layered perceptron, the generator is modelled as a differential function G , which maps the latent space ζ as $G(\zeta; \theta_g)$, while θ_g are the perceptron parameters. In terms of the discriminator, it is modelled as $D(x; \theta_d)$ with a scalar output, and x is the real input data [10]. The discriminator (D) and the generator (G) are trained concurrently. The discriminator’s task is to maximise the probability of correctly assigning labels to both the synthetic signal $G(\zeta)$ and the realistic signal x , while the generator’s task is to minimise the $\log(1 - D(G(\zeta)))$. This is the minimisation–maximisation of the value function $V(G, D)$ [10]:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{\zeta \sim p_{\zeta}(\zeta)} [\log(1 - D(G(\zeta)))] \tag{8}$$

Figure 3 demonstrates the alternate training of the generator and the discriminator. The discriminator receives synthetic examples \hat{X} produced by the generator from input noise and real examples (X). Afterwards, the discriminator makes predictions \hat{Y} , which is the probability score of how synthetic and how real each of the signals is. The predictions are compared using BCE loss with the desired labels for synthetic and real signals. This process leads to the update of θ_d , where d denotes the parameters for the discriminator. The generator is only updated by θ_g (g denotes the generator), which are the parameters resulting from the synthetic signals only. That is, the generator training is not affected by the real signals.

After cost computation, the gradient is propagated backwards, and the parameters of the generator θ_g are updated. As the alternate training occurs, models are trained in succession. Therefore, in the alternate training of GANs, it is important for both models to improve concurrently, and they should be kept at similar skill levels from the beginning of training. In terms of the GAN loss, this research used the BCE loss function (Equation (9)). The BCE loss measures the divergence between two probabilities; in the context of a GAN, the predicted probability distribution is the output of the discriminator, which indicates whether the input is real (1) or synthetic (0).

$$J(\theta) = -\frac{1}{m} \sum_{i=0}^m \underbrace{[y^{(i)} \log h(x^{(i)}, \theta)]}_{L1} + \underbrace{[(1 - y^{(i)}) \log(1 - h(x^{(i)}, \theta))]}_{L2} \tag{9}$$

h: prediction,
 X: features,
 θ : parameters,
 y: label,
 J: average patch loss.

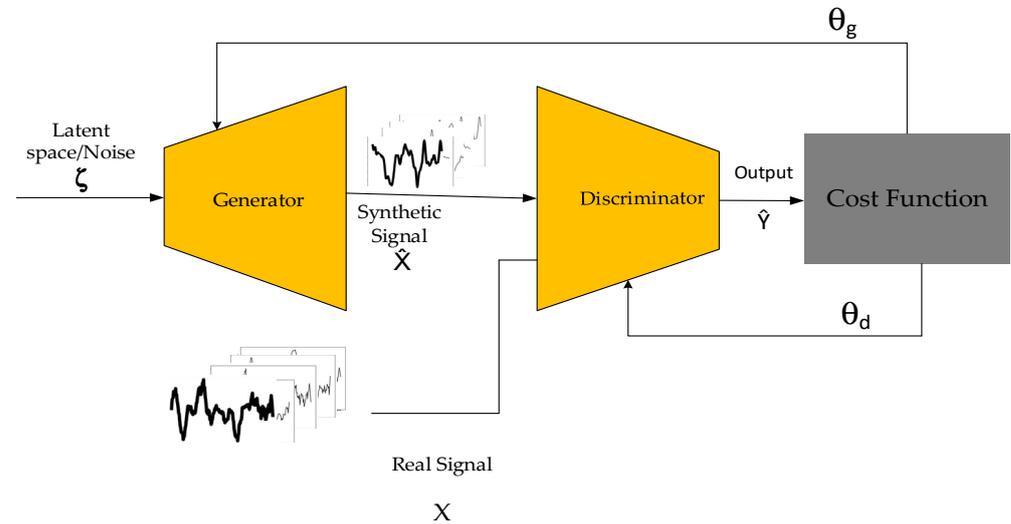


Figure 3. GAN components.

In Figure 4, the L1 part of the equation approaches zero when the real value is correctly detected, while the L2 component of the equation yields zero when a synthetic signal is successfully identified. If the model incorrectly classifies either part’s relevant label, a significant loss will result for both.

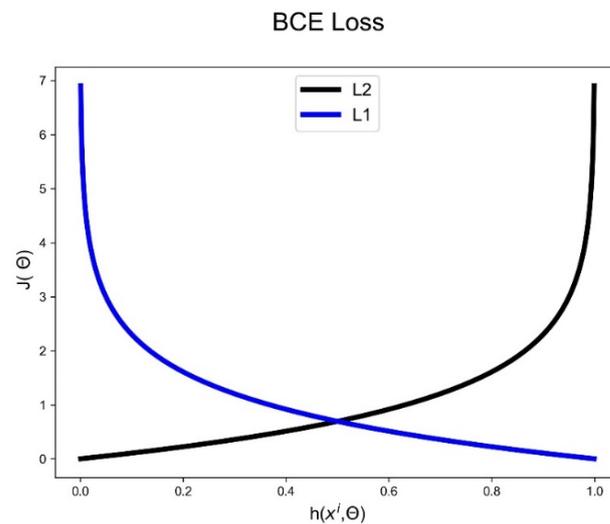


Figure 4. The graph of the two terms comprising the BCE loss function.

Proposed Architecture

The 1DDCGAN network was implemented in Python, while MATLAB was used to extract features from the sEMG signals to be classified during the evaluation phase. Latent noise is fed into the generator, which produces a synthetic sEMG signal. FFT, DTW and wavelets are extracted from the synthetic and real signals and then fed into the discriminator, which distinguishes between real and synthetic signals and updates the generator accordingly. The discriminator in this study used the DTW, FFT and wavelet techniques, which were adopted from previous work [19] but in a different context, which

employed GANs and style transfer to produce sEMG signals mimicking the tremor pattern of individuals with Parkinson’s disease. The present study did not incorporate the style transfer component.

The layers of the generator and the discriminator both used batch normalisation, dropout, 1D convolution and leaky ReLU. Batch normalisation ensures that the values input to each layer have constant scales and distributions across the batch, reduces internal covariate shifts and enhances training stability and convergence, using mini-batch statistics to normalise the data input to each layer [27]. Dropout is a regularisation approach, which is intended to enhance generalisation and lessen overfitting. The generator and discriminator networks use 1D GANs and convolutional layers to learn and extract meaningful features from the input data. Negative values can have a non-zero slope using the leaky ReLU activation function, preventing mode collapse.

- Generator

The generator network consists of six layers: convolution, batch normalisation, dropout and ReLU activation. Three of them have an up-sampling layer to increase the input dimensions. The six layers are connected to a dense layer and a Tanh activation at the end (Figure 5).



Figure 5. Generator architecture.

In the layers of the generator, up-sampling and convolution are used instead of deconvolution because deconvolution can be computationally expensive and introduce artefacts into the signal. Up-sampling is used to increase the size of the input by creating new data points between existing data points; in this paper, nearest neighbour interpolation was used.

The output of the generator contains a hyperbolic tangent (tanh) activation function—which prevents the generator from producing unrealistic or out-of-range samples—and increases GAN training stability and convergence, resulting in higher quality output.

- Discriminator

The input to the discriminator is the generator’s output and the participants’ original signals. The input is passed through the convolution layer, and the output is concatenated with the output of the same convolution layer but with input from the mini-batch discriminator, FFT, DTW and wavelet filters. In Figure 6, the discriminator consists of seven layers containing convolution, batch normalisation, leaky ReLU and dropout.

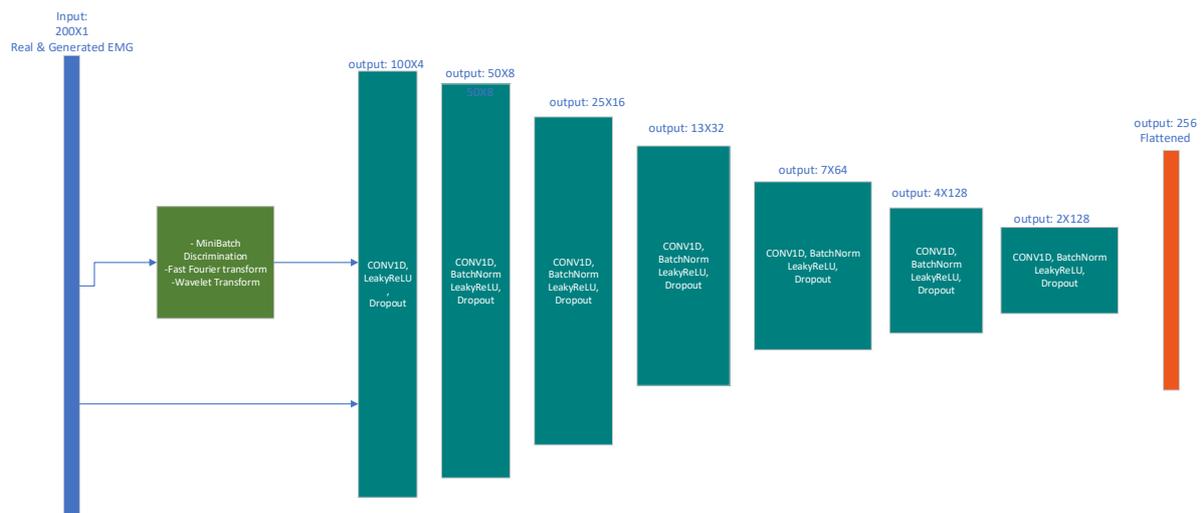


Figure 6. Discriminator architecture.

Salimans et al. (2016) [27] proposed several recommendations to improve GAN performance, including one-sided label smoothing to improve training stability and prevent mode collapse. This works by substituting hard labels (0/1) with soft labels (0.1/0.9) to train the discriminator.

By replacing the target labels with real examples with a slightly lower value (denoted as α) and synthetic signal targets (β), the discriminator model can be described as

$$D(x) = \frac{\alpha p_{\text{data}}(x) + \beta p_{\text{model}}(x)}{p_{\text{data}}(x) + p_{\text{model}}(x)} \quad (10)$$

When the p_{model} is large and the p_{data} are very small in value, the samples have no incentive to move closer to the data. Therefore, the positive labels (α) are only smoothed.

2.4. Evaluation

It is difficult to judge the quality of the images generated by GANs, let alone of those generated by sEMG signals. When it comes to images, humans can be employed to assess their quality. However, it is not feasible to have humans judge the quality of synthetic sEMG signals. This presents an obstacle to assessing the quality of generated sEMG signals and requires alternative methods, such as objective metrics or expert evaluation, to evaluate the effectiveness of the GAN. The features from Section 2.2.2 will be used for evaluation, since they offer more insightful data regarding the current window than the raw signal.

Salimans et al. [27] introduced the inception score (IS) and the Fréchet inception distance (FID). The IS is an objective metric for evaluating the generated images, while the FID is used to measure the distance between the features of the synthetic dataset and the realistic one.

Delaney et al. [28] introduced DTW as a metric for time series data, and it has been used successfully in Ref [17] to evaluate the model’s performance. Esteban et al. [14] quantitatively used sample likelihood and MMD to compare the synthetic ICU data with the original dataset. Hazra et al. [15] and Hernandez-Matamoros [29] used the RMS error, percentage RMS and other metrics to evaluate the generated dataset.

This study utilises the Mantel test and C2ST and classification techniques to evaluate the quality of the generated signals following the training process. Furthermore, the ML models were utilised to classify both the original signal and the signals generated through synthetic signal augmentation to investigate the impact of this augmentation technique. To the author’s knowledge, the application of the Mantel test and C2ST in this context has not been previously documented.

2.4.1. Mantel Test

The Mantel test is a statistical technique employed to assess the correlation between distance matrices via the Mantel test [30]. It compares two distance matrices by calculating the correlation among the lower or upper triangles of the symmetric distance matrices. The terminology “distance matrix” is utilised in this research to refer to the correlation matrix between the features of a provided signal. The Mantel test makes it easier to investigate whether the produced signal can maintain the inherent relationship among features in the original signal. The present study employed Pearson’s product–moment correlation coefficient with 1000 permutations.

According to Mantel’s definition [31], the Mantel test calculates a test statistic r_M based on two symmetric distance matrices, namely D_X and D_Y . r_M is defined as

$$r_M = \frac{1}{d-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{stand}(D_X)_{ij} \text{stand}(D_Y)_{ij} \quad (11)$$

$$d = \frac{n(n-1)}{2}$$

where “n” denotes the total number of rows and columns in all the distance matrices. The matrices $\text{stand}(D_X)$ and $\text{stand}(D_Y)$ consist of standardised distances within their respective upper triangles and are both classified as distance matrices.

2.4.2. Classification

One of the aspects determining the quality of the generated signal is its ability to help the classification model achieve a better classification result, which is one of the main reasons GANs and other signal synthesis models are used. This research used the LazyPredict library in Python, which contains 36 models divided into three categories: linear models, tree-based models and ensemble models. Five distinct sliding windows with varying increments were employed to extract features, and then, these features were used as inputs to the classifiers. The StandardScaler was employed for feature scaling and centring; it standardises the dataset by transforming it to have a mean of zero and a variance of one, resulting in enhanced model performance and facilitated convergence of the model [32]. The best results were obtained from ensemble models, in particular, the ADABOOST and bagging classifier. AdaBoost, short for adaptive boosting, is a machine-learning algorithm, which combines weak classifiers to create a strong classifier by giving more importance to misclassified instances. Bagging [33], short for bootstrap aggregating, is a machine-learning technique. Bagging classifiers use multiple independent classifiers trained on different subsets of data to improve prediction accuracy and reduce the impact of outliers or noisy data [34].

2.4.3. Augmentation Classification Improvement

Synthetic signals generated from GAN are added incrementally to assess its effect on the classification results of the model used and to examine whether the augmentation can improve the classification result. The increments are 25% of the original signal. Each increment is balanced in terms of the amount of data for each gesture compared to the original data. Support vector machine (SVM) was used for classification in this section. SVM is an established technique for pattern detection in myoelectric signals. Its main purpose is to find an n-dimensional hyperplane, which divides a set of input feature points into distinct classes. With this method, complex patterns may be detected [35].

2.4.4. Classifier Two-Sample Test

Two-sample tests are employed to ascertain the equality of two probability distributions and their similarity. C2ST can be utilised to accomplish the same objective, and a binary classifier can be employed to evaluate the quality of the produced distribution [36,37]. Osokin et al. [38] used classifier two samples to evaluate their GAN model, which was designed to produce synthetic medical images. Liu et al. [39] and Pacchiardi et al. [40] used C2ST for anomaly detection in videos.

Two samples are selected from distributions G and R, which represent the generated and reference distributions, respectively. The positive and negative labels are assigned to R and G. For the null hypothesis $P = R$ to hold true, the results should remain close to the chance level, indicating that the two distributions are equivalent. The mean of the individual features determines the accuracy of the test and converges towards a Gaussian null distribution. Lopez-Paz et al. [36] designed a dataset comprising $x_i, y_j \in \mathcal{X}$, where $i = 1, \dots, n$, and $j = 1, \dots, m$. To determine the equality of G and R, the construction of the dataset is as follows:

$$\mathcal{D} = \{(x_i, 0)\}_{i=1}^n \cup \{(y_j, 1)\}_{j=1}^m := \{(z_i, l_i)\}_{i=1}^{n+m} \quad (12)$$

Dataset \mathcal{D} is randomly shuffled and then divided into two distinct subsets, namely the training (\mathcal{D}_{tr}) and testing datasets (\mathcal{D}_{te}), which correspond to 80% and 20% of the dataset to avoid overfitting and obtain an accurate estimation of the classifier's performance.

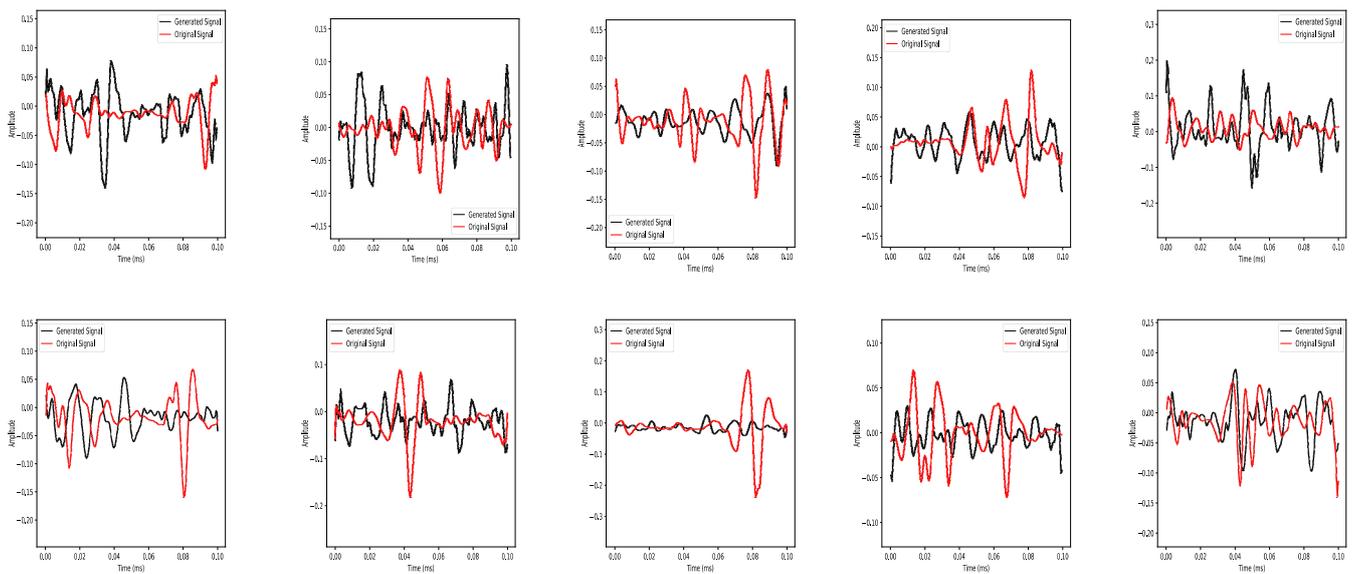
The training process of the binary classifier is $f : \mathcal{X} \rightarrow [0, 1]$ for \mathcal{D}_{tr} , while $f(z_i)$ is the probability distribution, $p(l = 1 | z_i)$.

$$\hat{t} = \frac{1}{n_{\text{te}}} \sum_{(z_i, l_i) \in \mathcal{D}_{\text{te}}} \mathbb{I} \left[\left(f(z_i) > \frac{1}{2} \right) = l_i \right] \quad (13)$$

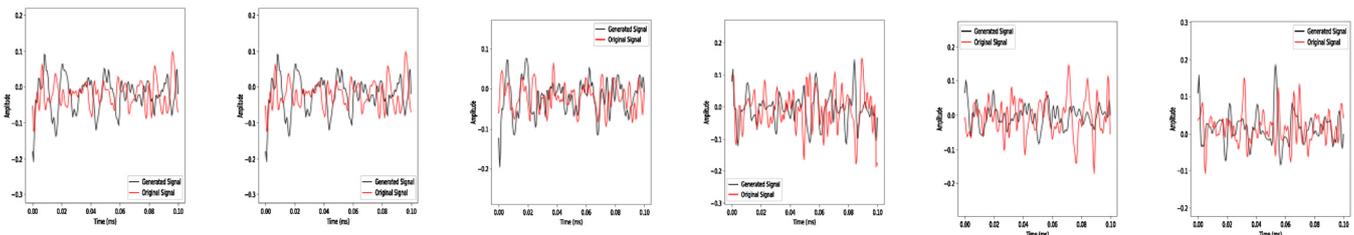
Throughout the GAN training process, the generator attempts to generate a probability distribution, which closely approximates that of the discriminator. Nevertheless, in actuality, the GAN will not generate an identical distribution or serve as a mere replica of the input. Lopez-Paz et al. [36] suggested using a margin classifier with a finite norm. Consequently, in the present investigation, the margin classifier with a finite norm implemented was logistic regression due to its ease of training and utilisation.

3. Results

Visually, the GAN attempts to produce signals in Figure 7 mimicking the original signals. However, a complete duplicate of the original signal must be avoided, as this would signify mode collapse. Mode collapse is a bad behaviour, which shows a lack of diversity in the samples, information loss and an inability to cover the complete target distribution. As a result, the quality of the generated signals will not be assessed based on the original sEMG signals. Instead, features from the time and time-frequency domains will be retrieved and analysed. The relationships between features in the reference and generated signals are examined in Section 3.1. We evaluate the categorisation precision of both sets of signals in Section 3.2. The improvement in classification accuracy attained through the augmentation of the generated signal is covered in Section 3.3. Last but not least, Section 3.4 compares the distributional similarity between the reference and generated signals using a classifier two-sample test. By using this structured technique, the generated signals can be thoroughly evaluated, along with their impact on classification accuracy and their statistical resemblance to the reference signals.



(a) Random samples from classes 1 to 10, respectively, from the public dataset.



(b) Random samples from classes 1 to 6, respectively, from the private dataset.

Figure 7. Random samples of reference signals and their corresponding generated signals of all classes of both datasets.

3.1. Mantel Test Analysis

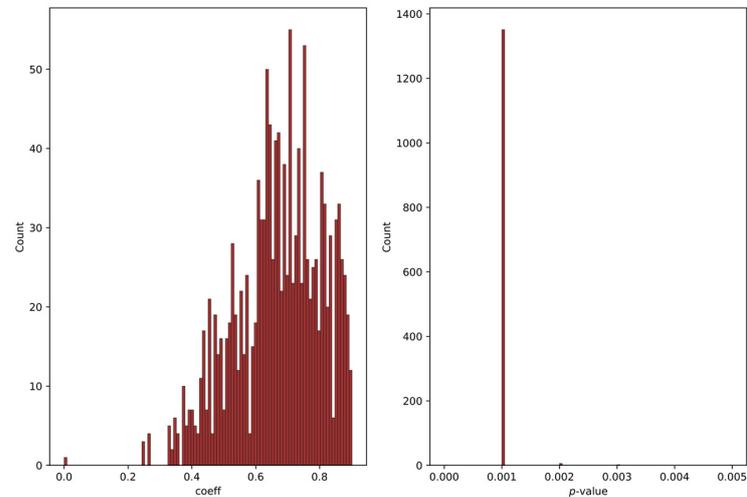
This visual method of examining the correlation matrix is laborious and impossible for a large amount of data. Therefore, the Mantel test is utilised, which describes the similarity between the two matrices using a correlation coefficient and a p -value. The p -value of the Mantel test is the probability of obtaining a correlation coefficient as extreme as the one observed, assuming that the null hypothesis of no correlation is true. A low p -value (typically < 0.05) indicates that the correlation is statistically significant.

Figure 8 represents the correlation coefficient between correlation matrices of the original and synthetic signals for both datasets and the corresponding p -value for both the public and private datasets. For both datasets, the vast majority of the correlation coefficient indicates a positive correlation with a p -value < 0.05 for the public dataset, indicating that the results are statistically significant. In terms of the private dataset, a minimal number of samples have a p -value > 0.05 , indicating that very few of the results are statistically insignificant.

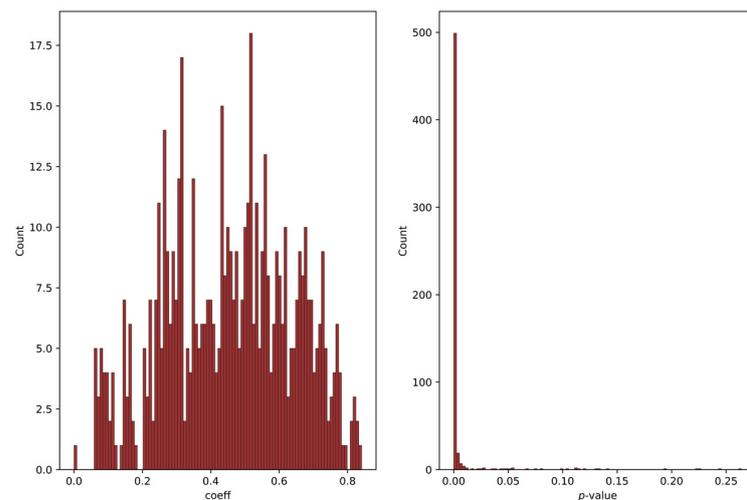
For the private dataset, 9% of the samples showed a very weak correlation; 31% of its samples showed a weak positive correlation; 36% showed a medium correlation; 23% showed a strong correlation; and 1% showed a very strong correlation. In terms of the public dataset, 20% of the samples showed a very strong correlation; 53% showed a strong correlation; 23% showed a medium correlation; and the rest were in a weak correlation range with only 4%.

Both the private and public datasets show a positive correlation. All p -values for the public dataset are less than 0.05; for the private dataset, 96.3% have values less than 0.05. This indicates that the null hypothesis is rejected for most of the samples in both datasets,

and the results are statistically significant. Hence, it can be deduced that the GAN model used has the capability of producing samples, which exhibit a resemblance of the complex interrelationships among the diverse features of the original signal.



(a) Public dataset.

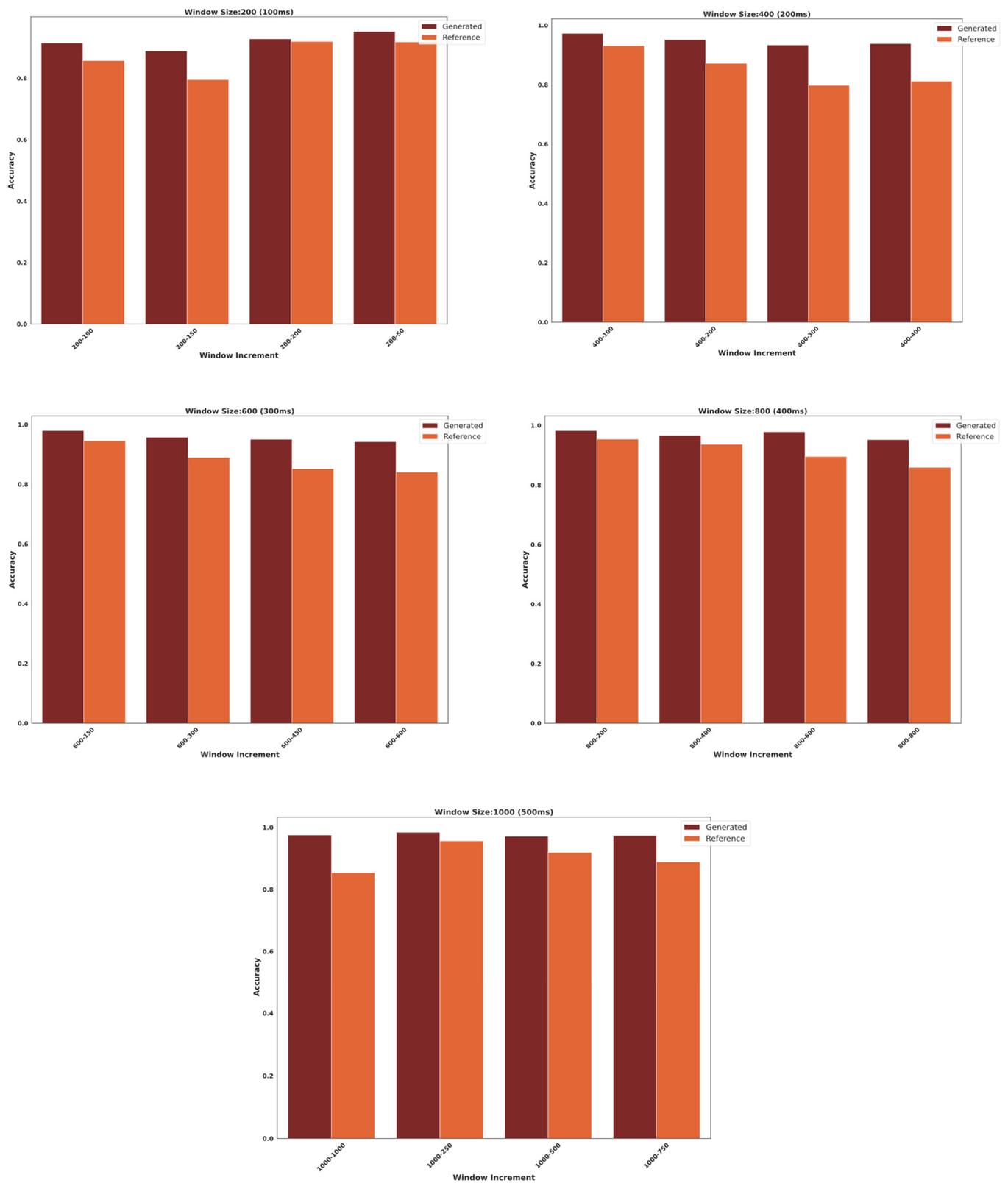


(b) Private dataset.

Figure 8. Mantel's correlation coefficient and p -value for public and private datasets. The horizontal axis denotes the numerical value of the p -value, while the vertical axis indicates the frequency of its occurrence.

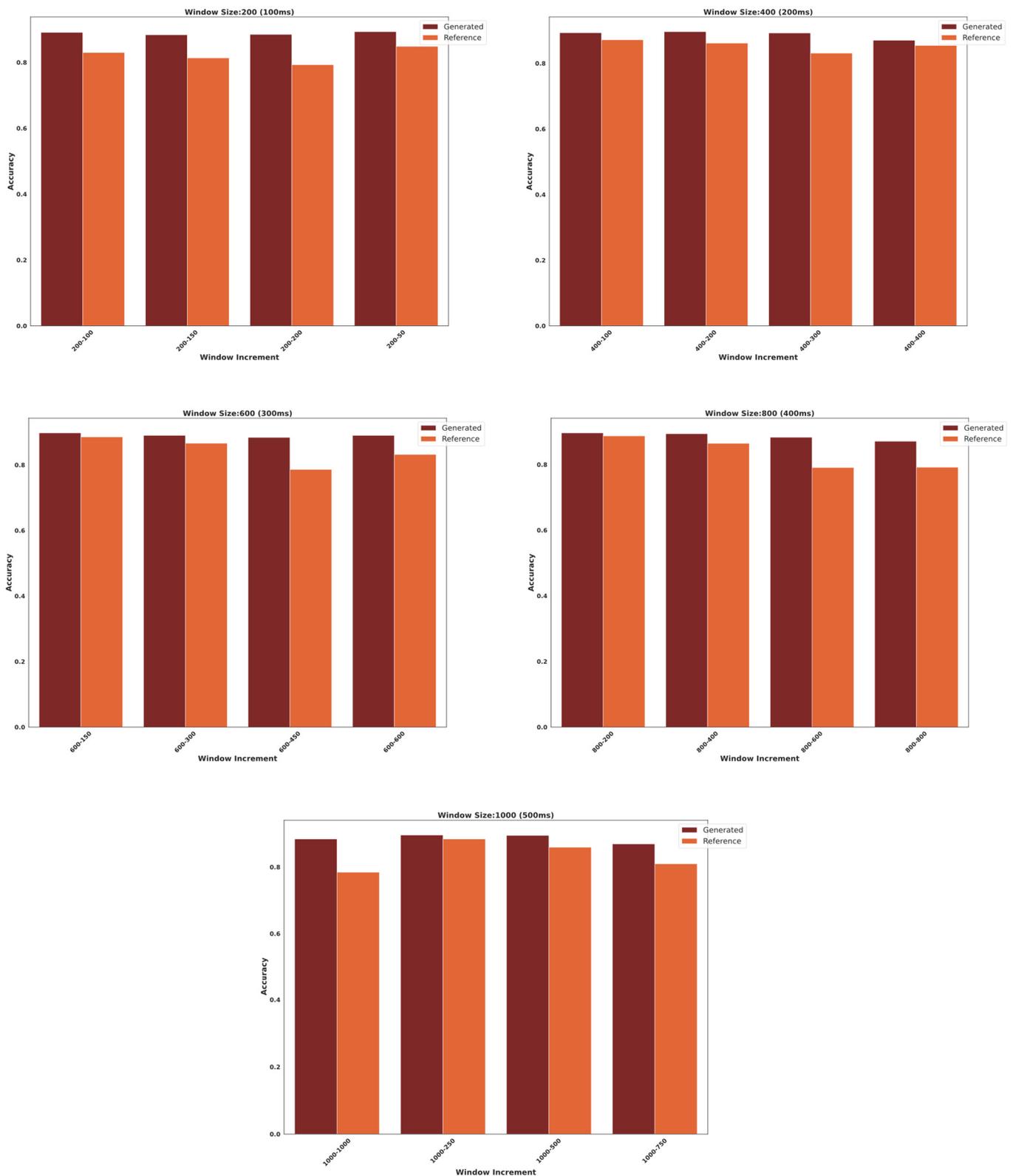
3.2. Classification

In order to further analyse the synthetic and real data, both synthetic and reference data were classified on their own using the LazyPredict library. The models were trained on 80% of the dataset to avoid overfitting and obtain an accurate assessment of their performance. The remaining 20% of the dataset was used to evaluate the accuracy of the models, as illustrated in Figure 9. The best results were achieved in ensemble models AdaBoost and bagging classifier.



(a) Private classification of real and synthetic signals, each on its own.

Figure 9. Cont.



(b) Public classification of real and synthetic signals, each on its own.

Figure 9. Accuracy results of private dataset (a) and public dataset (b).

The classification models were trained to predict the correct gestures per subject, using the features extracted from each window for that subject. The windows were applied in all trials and only in the gesture part of the trial and not the rest part. The gestures were 3 s

long and repeated 15 times. The classification results of each subject for all windows were then averaged across both datasets and presented in Figure 9.

The generated sample of the private dataset achieved its highest classification accuracy of 98.4% using a window size of 1000 and an increment of 250 data points versus 95.6% for the original signal for the same window and increment with a difference of 2.8%. Overall, the average difference between the generated and reference samples was 6.9%.

In terms of the public dataset, the highest classification accuracy achieved was 89.6% for the sample generated with a window size of 1000 and an increment of 250 data points as well versus 88.4% accuracy for the original signal of the same window and increment with a difference of 1.2%. The average difference among all classified samples was 8.9%.

For both datasets (Figure 9a,b), the higher classification results for the synthetic signals can be attributed to the fact that the synthetic data generated from GAN are more informative and easier to classify than the original data because the intraclass distance is small and the interclass distance is large compared to that of the original signal. A small intraclass distance makes it easier for the classifier to find patterns within each class, and a substantial interclass distance makes it easier for the classifier to distinguish between different classes. This indicates that the GANs are able to learn the underlying distribution of the data and generate samples, which are more representative of the real world, and that the machine-learning model is able to generalise to the target distribution.

3.3. Augmentation Classification Performance Analysis

Signals were divided into five different window sizes with lengths of 200, 400, 600, 800 and 1000 data points, which corresponded to 50, 100, 150, 200 and 250 milliseconds, respectively, for the public dataset and 100, 200, 300, 400, 500 for the private dataset, as seen in Table 1. The window increments used were 25%, 50%, 75% and 100% of the specific window size. The features from these windows were extracted and used in this classification. The datasets were divided into 80% for training and 20% for testing to obtain an accurate assessment of the model performance; the accuracy results recorded were those of the testing dataset. The testing dataset comprised a balanced distribution of both generated and original signals. Specifically, the percentage of generated signals in the testing dataset was consistent with the percentage of generated signals present in the entire mixed dataset. Windows were applied in all trials exclusively for the gesture segment, excluding the rest portion. The length of a gesture was 3 s, as mentioned in Figure 2, and it was repeated 15 times.

Table 1. The number of data points for both public and private datasets and the corresponding lengths of these data points. The datasets have different frequencies: 4000 Hz for the public and 2000 Hz for the private dataset.

Number of Data Points	Public Dataset (ms)	Private Dataset (ms)
200	50	100
400	100	200
600	150	300
800	200	400
1000	250	500

In order to study the effect of the generated synthetic signals on the performance of a classifier, 25%, 50%, 75% and 100% of the synthetic signals are added to the classification dataset, and the classification results are plotted in Figure 10 below.

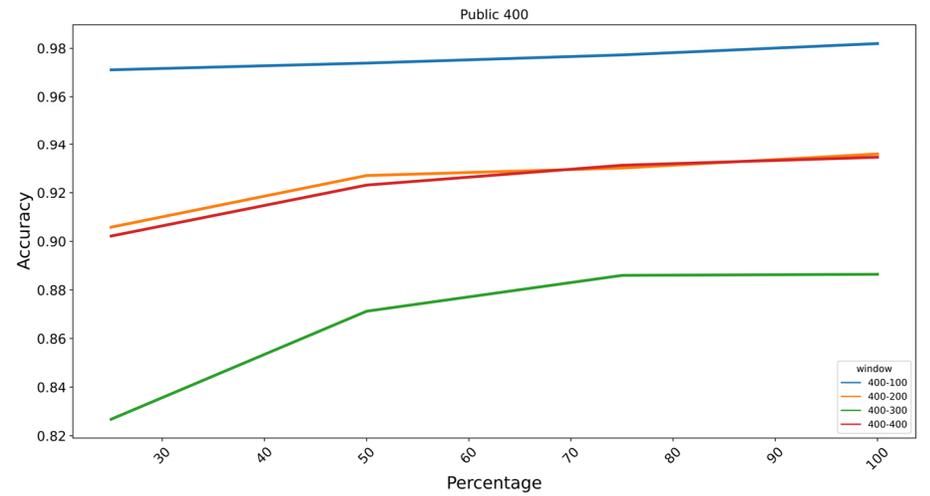
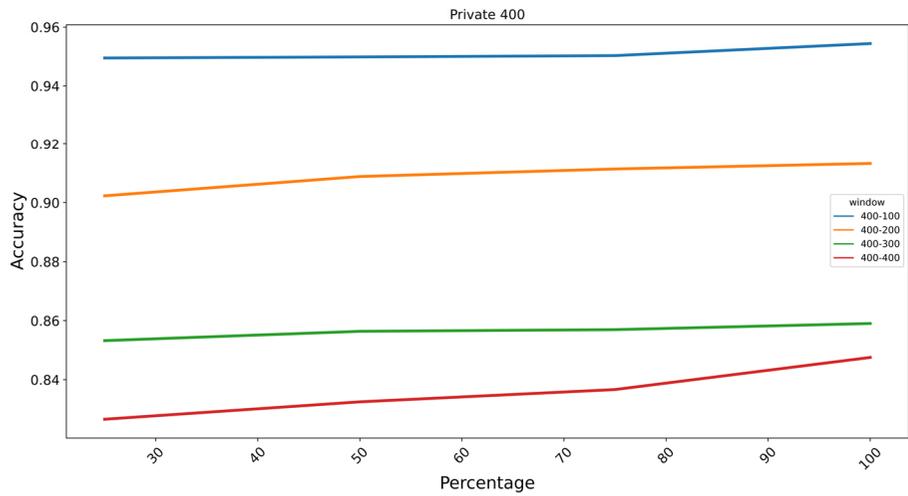
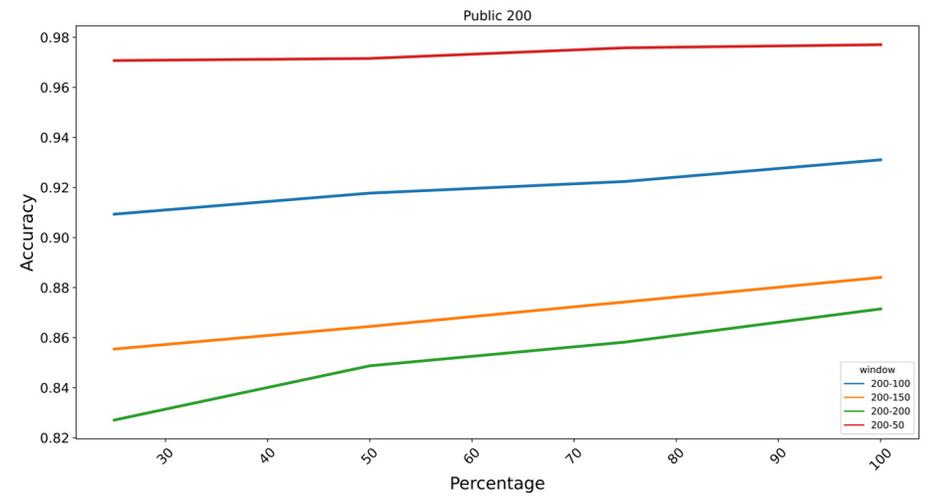
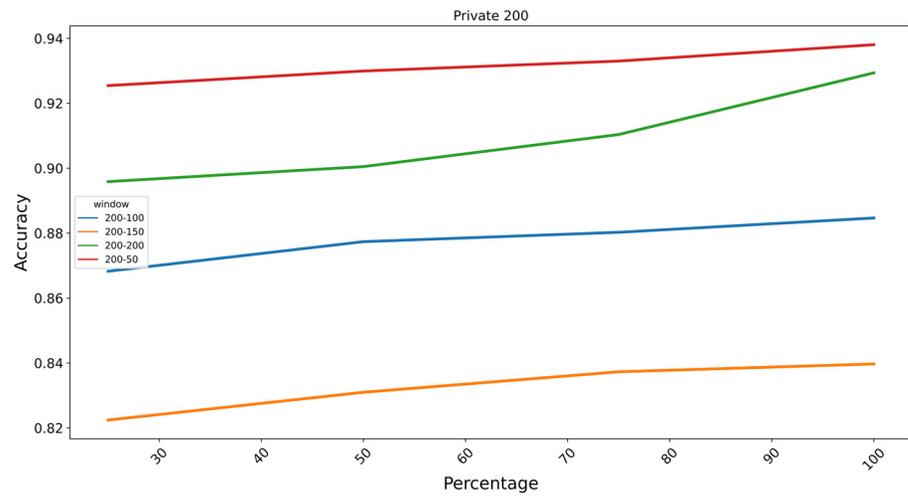


Figure 10. Cont.

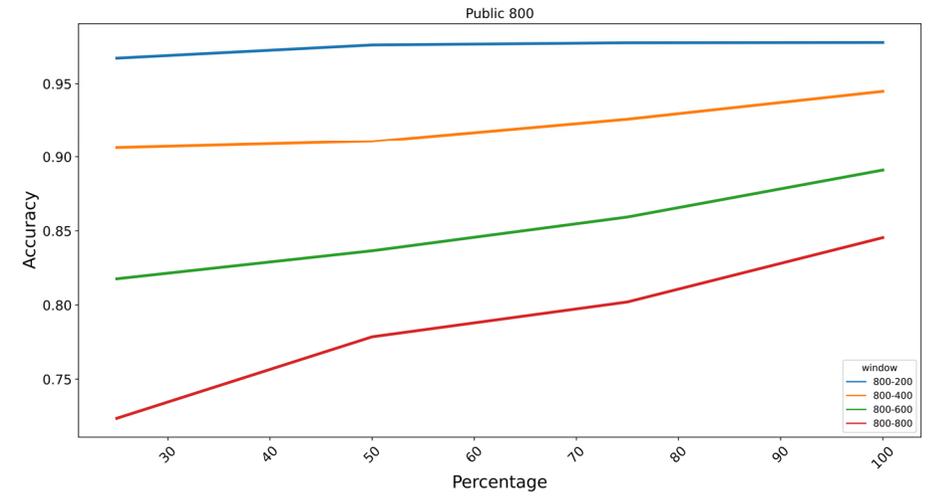
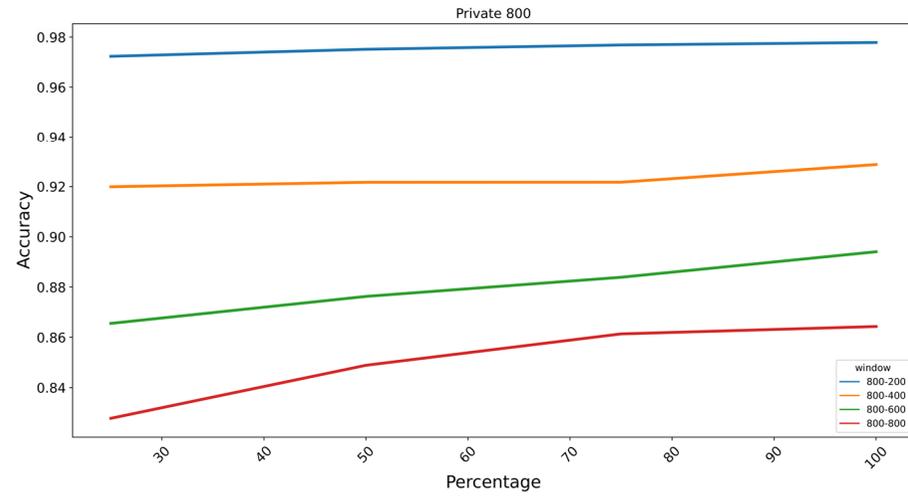
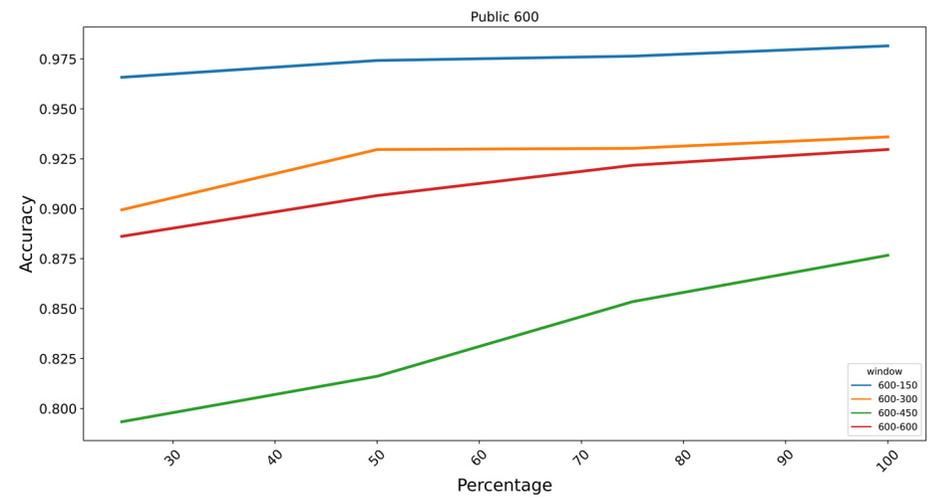
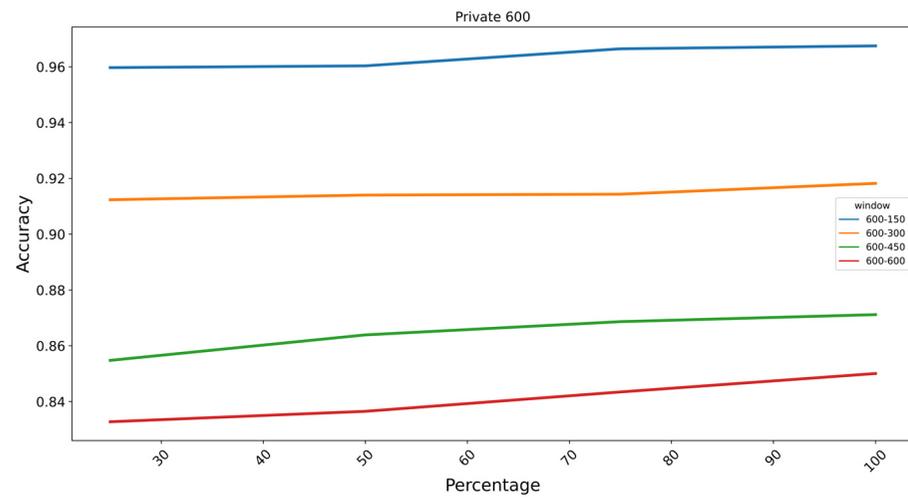


Figure 10. Cont.

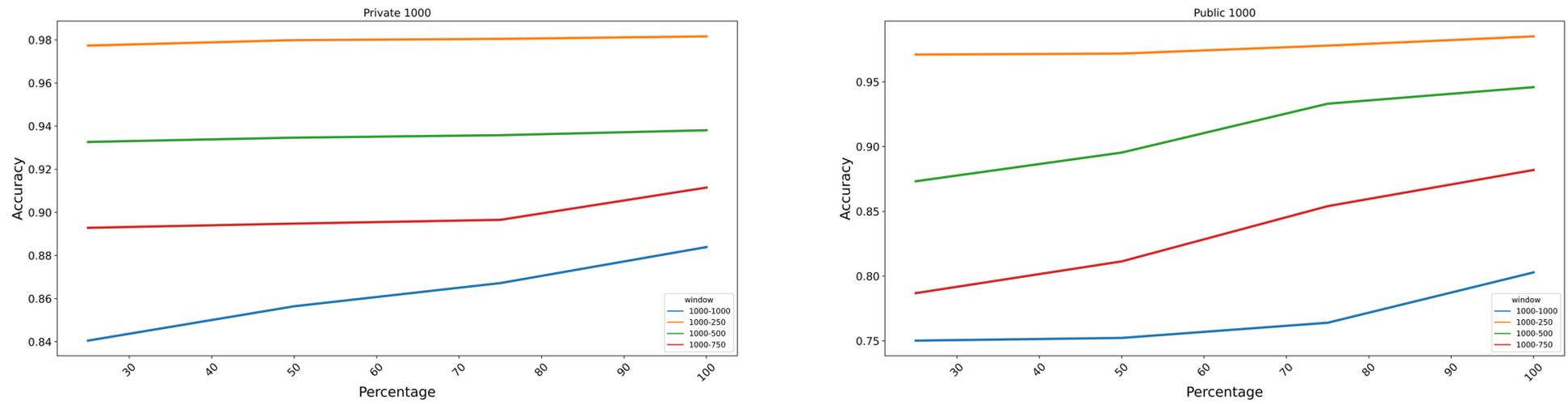
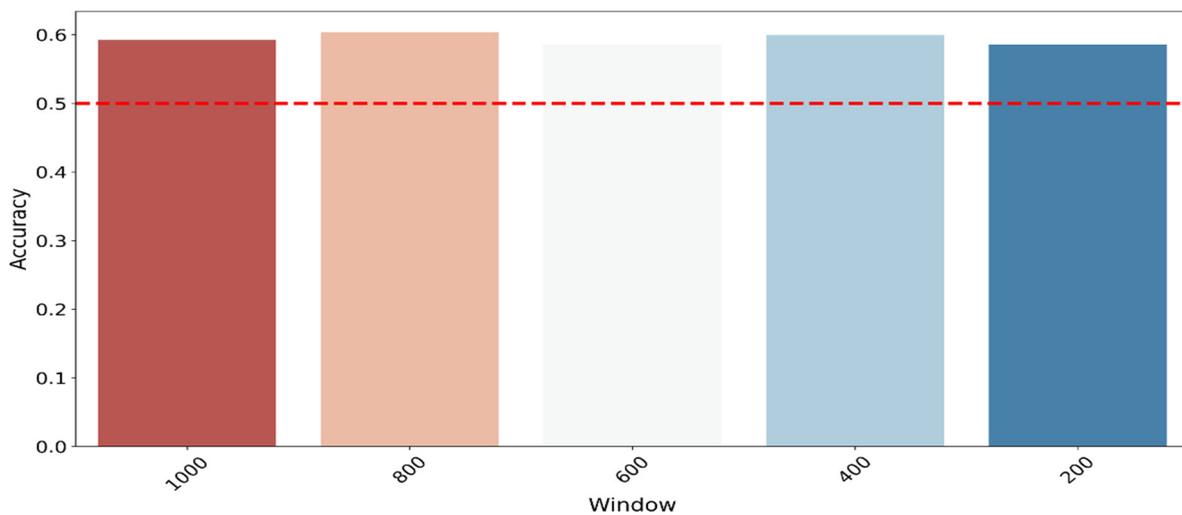


Figure 10. The figure represents the classification results of five different windows with four increments representing 25%, 50%, 75% and 100% of each window for both public and private datasets. The X-axis represents the percentage of synthetic signals added to the original signal during classification.

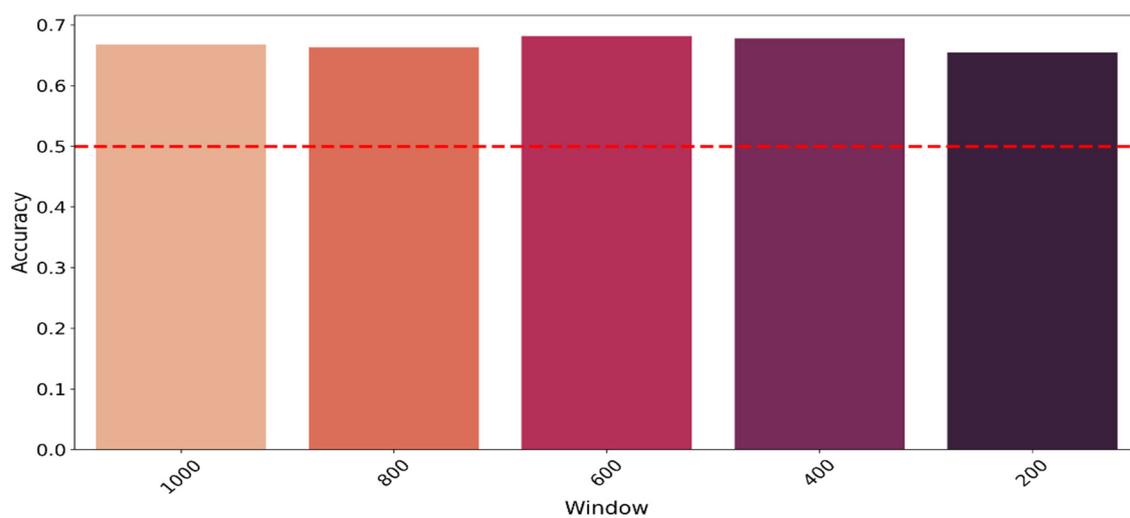
In Figure 10, it can be observed that the window increment of 25% of all windows achieved the highest results in both datasets. The window with the highest classification result was the one with a length of 1000 data points and an increment of 250 data points for both public and private datasets, with a classification accuracy of 98.5% and 98.1% using the SVM classifier, respectively. The best classification result corresponded—for both datasets—to 100% addition of synthetic data. The addition of generated samples to the original dataset improved the classification results by an average of 1.21% and 5% for the private and public datasets, respectively.

3.4. Classifier Two-Sample Test

Logistic regression is used for classification to investigate the similarity of two probability distributions, and it is widely used and easy to train. The results for both datasets are presented in Figure 11a,b. The private dataset's average accuracy for each window is presented in Figure 11a; the average accuracy of all windows is 59%. In terms of the public dataset, the average accuracy can be found in Figure 11b, and the average of all window accuracies is 66%.



(a) Private dataset C2ST results.



(b) Public dataset C2ST results.

Figure 11. Private (a) and public (b) dataset C2ST results.

While the average in both datasets is higher than the threshold of 50%, it is still close to it, especially for the private dataset. This indicates that while the generated data are not perfectly similar to those of the reference signal, they are similar to some extent.

4. Discussion

To the best of the author's knowledge, the use of 1DDCGAN for the generation of sEMG signals corresponding to hand gestures has not been previously explored. Additionally, two novel metrics were introduced—namely the Mantel test and C2ST—for evaluating the quality of the generated signals. These metrics have not been utilised in this context before.

The majority of the Mantel test results indicated a medium-to-strong correlation with a p -value less than 0.05, indicating that the correlation between two matrices is statistically significant at a 95% confidence level, meaning that there is strong evidence to reject the null hypothesis of no correlation between the matrices. The Mantel test results suggest that the proposed GAN successfully captured the inherent relationship between the extracted features in the real signal and replicated it in the generated signal. This finding demonstrates the effectiveness of the GAN model used in accurately reflecting the underlying patterns of the real data and in generating high-quality signals.

In Section 2.3, the synthetic and original datasets were independently subjected to classification using multiple classifiers. The classification results were found to be the best when using a data window size of 1000 and an increment of 250. In this configuration, the generated data achieved a classification accuracy of 98.4%, while the original data achieved a lower accuracy of 95.6%. Consequently, the generated data outperformed the original data by about 2.8% in classification. The results indicate that the generated data consistently exhibited higher accuracy compared to the original data across all signal classifications. This suggests that the proposed 1DDCGAN model effectively produced signals characterised by small intraclass distance and large interclass distance, thereby facilitating easier classification. The observed outcome is consistent with the expected output of a GAN, suggesting its ability to generate signals, which show similarity within a given class while also demonstrating sufficient dissimilarity from other classes, resulting in notable classification performance.

Regarding the augmentation test in Section 3.3, the synthesised data were augmented by adding increments of 25% of the original signal's length until reaching a ratio of 1:1 with the original signal. Subsequently, the augmented data were classified using an SVM classifier. The results indicate that the augmentation technique led to enhanced classification outcomes for all of the windows. It is noteworthy to mention that the best outcomes were also attained when using a window size of 1000, with increments of 250 data points, which aligns with the findings of the classification analysis discussed in Section 3.2. The overall improvement varied between 1.21% and 5% for both the public and private datasets, with a maximum improvement of 12.2%.

In terms of the C2ST results, the perfect result would be as near to the chance level as possible, as this would indicate that the two samples are indistinguishable. The results obtained for the private dataset and the public dataset were 58% and 66%, respectively. These results suggest that there is a sufficient degree of similarity between the generated signals and the original signals, indicating that the proposed model is capable of producing signals, which closely resemble the original data. The classification results show that the augmentation of the generated dataset can enhance the classification result.

In summary, the evaluation metrics results demonstrated that the proposed model outperformed the real data in classification while effectively reflecting the underlying patterns of the real data. The results of the classification were improved once the generated dataset was augmented, showing that GAN-generated samples can offer more varied and representative examples for the classification model to learn from.

5. Conclusions

The approach put forth in this study involved using 1DDCGAN to generate signals. In addition, multiple evaluation metrics were used to assess the generated data. The metrics used were the Mantel test, classification, augmentation test and C2ST. The Mantel test and C2ST are novel in this context. The results indicated that the produced signals were able to capture the complex relationship among features and showed a smaller interclass distance and a large intraclass distance. The ability of 1DDCGAN to produce signals varied, and representative examples were shown in the augmentation test, where the augmentation of the synthetic signals achieved an enhancement in the classification results of about 1.21–5%. Furthermore, the C2ST results showed the ability of 1DDCGAN to produce signals similar to the original ones.

The findings indicate that the utilisation of 1DDCGAN demonstrated efficacy in generating signals suitable for augmenting signals in AI-controlled prostheses. This augmentation process has the potential to reduce the duration of laboratory training required for amputees to train the classification model of the prosthesis. Consequently, this approach offers benefits, such as increased comfort and time savings for amputees, as well as improved classification outcomes. Furthermore, the 1DDCGAN technique can also be employed to supplement datasets, which contain missing or corrupted data. Additionally, the Mantel test outcomes revealed that the generated signals exhibited a comparable interrelationship among their features compared to the original signals. This finding is significant, as it indicates that the 1DDCGAN model has the ability to generate signals with similar characteristics, making it a dependable tool for analysis and research purposes. Moreover, the model ensures the privacy of signal owners, thereby enabling researchers to access valuable data, which can contribute to advancements in the field of sEMG signal classification for gesture recognition.

Author Contributions: M.A.G.: Conceptualisation, Data curation, Formal analysis, Investigation, Writing—original draft; W.H.: Conceptualisation; D.J.: Data curation, Writing—review and editing; N.F.: Data curation; B.Z.: Data curation; Z.L.: Data curation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the University Innovation Team of Liaoning Province (LT2014006).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study will be made available upon request to the authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Toledo-Pérez, D.C.; Rodríguez-Reséndiz, J.; Gómez-Loenzo, R.A. A Study of Computing Zero Crossing Methods and an Improved Proposal for EMG Signals. *IEEE Access* **2020**, *8*, 8783–8790. [[CrossRef](#)]
2. Aviles, M.; Sánchez-Reyes, L.-M.; Fuentes-Aguilar, R.Q.; Toledo-Pérez, D.C.; Rodríguez-Reséndiz, J. A Novel Methodology for Classifying EMG Movements Based on SVM and Genetic Algorithms. *Micromachines* **2022**, *13*, 2108. [[CrossRef](#)] [[PubMed](#)]
3. Aviles, M.; Rodríguez-Reséndiz, J.; Ibrahimi, D. Optimizing EMG Classification through Metaheuristic Algorithms. *Technologies* **2023**, *11*, 87. [[CrossRef](#)]
4. Malešević, N.; Olsson, A.; Sager, P.; Andersson, E.; Cipriani, C.; Controzzi, M.; Björkman, A.; Antfolk, C. A database of high-density surface electromyogram signals comprising 65 isometric hand gestures. *Sci. Data* **2021**, *8*, 63. [[CrossRef](#)] [[PubMed](#)]
5. Montazerin, M.; Rahimian, E.; Naderkhani, F.; Atashzar, S.F.; Yanushkevich, S.; Mohammadi, A. Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals. *Sci. Rep.* **2023**, *13*, 11000. [[CrossRef](#)] [[PubMed](#)]
6. Brophy, E.; Wang, Z.; She, Q.; Ward, T. Generative Adversarial Networks in Time Series: A Systematic Literature Review. *ACM Comput. Surv.* **2023**, *55*, 199. [[CrossRef](#)]
7. Kaczmarek, P.; Mańkowski, T.; Tomczyński, J. putEMG—A Surface Electromyography Hand Gesture Recognition Dataset. *Sensors* **2019**, *19*, 3548. [[CrossRef](#)] [[PubMed](#)]
8. Zhang, X.; Ma, M. Research on sEMG Feature Generation and Classification Performance Based on EBGAN. *Electronics* **2023**, *12*, 1040. [[CrossRef](#)]

9. Korosec, D.; Martinez, C.; Zazula, D. Parametric modelling of EMG signals. In Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Amsterdam, The Netherlands, 31 October–3 November 1996; Volume 1474, pp. 1470–1471.
10. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Networks. *arXiv* **2014**, arXiv:1406.2661. [[CrossRef](#)]
11. Kataoka, Y.; Matsubara, T.; Uehara, K. Image generation using generative adversarial networks and attention mechanism. In Proceedings of the 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), Okayama, Japan, 26–29 June 2016; pp. 1–6.
12. De Rosa, G.H.; Papa, J.P. A survey on text generation using generative adversarial networks. *Pattern Recognit.* **2021**, *119*, 108098. [[CrossRef](#)]
13. Wali, A.; Alamgir, Z.; Karim, S.; Fawaz, A.; Ali, M.B.; Adan, M.; Mujtaba, M. Generative adversarial networks for speech processing: A review. *Comput. Speech Lang.* **2022**, *72*, 101308. [[CrossRef](#)]
14. Esteban, C.; Hyland, S.L.; Rättsch, G. Real-valued (Medical) Time Series Generation with Recurrent Conditional GANs. *arXiv* **2017**, arXiv:1706.02633.
15. Hazra, D.; Byun, Y.-C. SynSigGAN: Generative Adversarial Networks for Synthetic Biomedical Signal Generation. *Biology* **2020**, *9*, 441. [[CrossRef](#)] [[PubMed](#)]
16. Beaulieu-Jones, B.K.; Wu, Z.S.; Williams, C.; Lee, R.; Bhavnani, S.P.; Byrd, J.B.; Greene, C.S. Privacy-Preserving Generative Deep Neural Networks Support Clinical Data Sharing. *Circ. Cardiovasc. Qual. Outcomes* **2019**, *12*, e005122. [[CrossRef](#)] [[PubMed](#)]
17. Anicet Zanini, R.; Luna Colombini, E. Parkinson’s Disease EMG Data Augmentation and Simulation with DCGANs and Style Transfer. *Sensors* **2020**, *20*, 2605. [[CrossRef](#)] [[PubMed](#)]
18. Campbell, E.; Cameron, J.A.D.; Scheme, E. Feasibility of Data-driven EMG Signal Generation using a1 Deep Generative Model. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; Volume 2020, pp. 3755–3758. [[CrossRef](#)]
19. Khushaba, R.N.; Kodagoda, S.; Takruri, M.; Dissanayake, G. Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals. *Expert Syst. Appl.* **2012**, *39*, 10731–10738. [[CrossRef](#)]
20. Kim, N.; Lim, T.; Song, K.; Yang, S.; Lee, J. Stretchable Multichannel Electromyography Sensor Array Covering Large Area for Controlling Home Electronics with Distinguishable Signals from Multiple Muscles. *ACS Appl. Mater. Interfaces* **2016**, *8*, 21070–21076. [[CrossRef](#)]
21. Ozdemir, M.A.; Kisa, D.H.; Guren, O.; Akan, A. Dataset for multi-channel surface electromyography (sEMG) signals of hand gestures. *Data Brief* **2022**, *41*, 107921. [[CrossRef](#)]
22. Krasoulis, A.; Vijayakumar, S.; Nazarpour, K. Multi-Grip Classification-Based Prosthesis Control with Two EMG-IMU Sensors. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 508–518. [[CrossRef](#)]
23. Atzori, M.; Müller, H. PaWFE: Fast Signal Feature Extraction Using Parallel Time Windows. *Front. Neurobot.* **2019**, *13*, 74. [[CrossRef](#)]
24. Hudgins, B.; Parker, P.; Scott, R.N. A new strategy for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.* **1993**, *40*, 82–94. [[CrossRef](#)] [[PubMed](#)]
25. Singla, R.; Chambayil, B.; Khosla, A.; Santosh, J. Comparison of SVM and ANN for classification of eye events in EEG. *J. Biomed. Sci. Eng.* **2011**, *04*, 62–69. [[CrossRef](#)]
26. Luca, C.J.D. The Use of Surface Electromyography in Biomechanics. *J. Appl. Biomech.* **1997**, *13*, 135–163. [[CrossRef](#)]
27. Salimans, T.; Goodfellow, I.; Zaremba, W.; Cheung, V.; Radford, A.; Chen, X. Improved Techniques for Training GANs. *arXiv* **2016**, arXiv:1606.03498.
28. Delaney, A.M.; Brophy, E.; Ward, T.E. Synthesis of Realistic ECG using Generative Adversarial Networks. *arXiv* **2019**, arXiv:1909.09150.
29. Hernandez-Matamoros, A.; Fujita, H.; Meana, H. A novel approach to create synthetic biomedical signals using BiRNN. *Inf. Sci.* **2020**, *541*, 218–241. [[CrossRef](#)]
30. Mantel, N. The detection of disease clustering and a generalised regression approach. *Cancer Res.* **1967**, *27*, 209–220.
31. Legendre, P.; Legendre, L. *Numerical Ecology*; Elsevier Science: Amsterdam, The Netherlands, 2012.
32. Anysz, H.; Zbiciak, A.; Ibadov, N. The Influence of Input Data Standardization Method on Prediction Accuracy of Artificial Neural Networks. *Procedia Eng.* **2016**, *153*, 66–70. [[CrossRef](#)]
33. Schapire, R.E. Explaining adaboost. In *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 37–52.
34. Skurichina, M.; Duin, R.P. Bagging for linear classifiers. *Pattern Recognit.* **1998**, *31*, 909–930. [[CrossRef](#)]
35. Oskoei, M.A.; Hu, H. Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 1956–1965. [[CrossRef](#)]
36. Lopez-Paz, D.; Oquab, M. Revisiting Classifier Two-Sample Tests. *arXiv* **2016**, arXiv:1610.06545.
37. Gretton, A.; Borgwardt, K.M.; Rasch, M.J.; Scholkopf, B.; Smola, A. A Kernel Method for the Two-Sample-Problem. In Proceedings of the 30th Conference on Neural Information Processing Systems (NIPS), Barcelona, Spain, 5–10 December 2006.
38. Osokin, A.; Chessel, A.; Carazo Salas, R.E.; Vaggi, F. GANs for biological image synthesis. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2233–2242.

39. Liu, Y.; Li, C.-L.; Póczos, B. Classifier Two-Sample Test for Video Anomaly Detections. In Proceedings of the British Machine Vision Conference, Newcastle, UK, 3–6 September 2018.
40. Pacchiardi, L.; Dutta, R. Likelihood-Free Inference with Generative Neural Networks via Scoring Rule Minimization. *arXiv* **2022**, arXiv:2205.15784.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.