Article

Flexible Covariance Matrix Decomposition Method for Data Augmentation and Its Application to Brainwave Signals

Hoirim Lee, Wonseok Yang [ID] and Woochul Nam *[ID]

Department of Mechanical Engineering, Chung-Ang University, Seoul 06974, Korea; hiu03091@cau.ac.kr (H.L.); yws1108@gmail.com (W.Y.)
* Correspondence: wcnam@cau.ac.kr; Tel.: +82-2-820-5270

Abstract: The acquisition of a large-volume brainwave database is challenging because of the stressful experiments that are required; however, data synthesis techniques can be used to address this limitation. Covariance matrix decomposition (CMD), a widely used data synthesis approach, generates artificial data using the correlation between features and random noise. However, previous CMD methods constrain the stochastic characteristics of artificial datasets because the random noise used follows a standard distribution. Therefore, this study has improved the performance of CMD by releasing such constraints. Specifically, a generalized normal distribution (GND) was used as it can alter the kurtosis and skewness of the random noise, affecting the distribution of the artificial data. For the validation of GND performance, a motor imagery brainwave classification was conducted on the artificial dataset generated by GND. The GND-based data synthesis increased the classification accuracy obtained with the original data by approximately 8%.

Keywords: brainwave signal; data augmentation; covariance matrix; generalized normal distribution

1. Introduction

Brain–machine interface (BMI) technology is aimed at acquiring and analyzing brainwave signals for device control [1,2]. BMI technology is used in fields [3] such as medicine, healthcare [4], education, and entertainment [5]. For example, patients can drive their wheelchairs [6], operate prosthetic hands [7], and control computer keyboards using their brainwave signals. Therefore, BMI can be used to address mobility inconveniences and speech difficulties caused by disabilities.

The quantity and quality of the data available for classification are important factors for the accurate estimation of a user's intention [8]. However, most BMI systems suffer from a lack of data because the acquisition of bio-signals requires stressful experiments. Various data augmentation techniques have been developed to address this issue, including noise addition methods, generative adversarial networks (GANs), and covariance matrix decomposition (CMD) methods. Noise addition methods generate artificial datasets by adding noise vectors to the original data [9–12]. For example, Salama et al. [11] added a Gaussian noise signal (with zero mean and unit variance) to electroencephalogram (EEG) signals to generate artificial EEG signals. However, this method adds simple noise without considering the stochastic characteristics of the bio-signal of interest. Therefore, a quality artificial dataset is not guaranteed when using this method.

GANs automatically learn the properties of the target bio-signal by using competitive networks (i.e., generator and discriminator) [13–16]. Luo et al. [15] suggested a conditional Wasserstein GAN for EEG data augmentation to improve the accuracy of emotion recognition. Zhang and Liu [16] used a conditional deep convolutional GAN method to generate artificial EEG data. However, GANs require a long training time and a large number of data samples [17]. Therefore, when only a small number of bio-signal samples are available, a GAN cannot generate high-quality artificial data.

CMD methods have been widely used to create stochastic signals because they consider the correlation between features [18–21]. CMD does not require complicated training; thus, its calculation time is very short. Furthermore, CMD provides high-quality data without an extremely large database. Owing to these advantages, CMD is an adequate data augmentation technique for bio-signals. CMD-generated artificial datasets increase the classification accuracy for brainwave, electromyography, and electrocardiography signals [21,22]. Therefore, in this study, CMD was used to generate artificial brainwave signals.

Accordingly, this study aimed to develop a more flexible CMD model than previous CMD models. CMD requires random noise to synthesize the artificial data. To preserve the correlations observed in the original data, the mean of the random noise should be zero, and its variance must be uniform. However, previous models impose another restriction on this random noise; they use only a standard normal distribution, even though this restriction is not related to correlation preservation. Thus, this study focused on releasing this restriction to provide greater flexibility to the CMD. The proposed model modifies the skewness and kurtosis of random noise by using a generalized normal distribution (GND). Then, the effects of skewness and kurtosis on accuracy were investigated for brainwave signals.

The remainder of this paper is organized as follows. Section 2 describes the motor imagery brainwave dataset used in this study and provides a detailed description of the proposed CMD method. Section 3 describes the artificial brainwave signals generated by the proposed method. The classification accuracies over different values of GND skewness and kurtosis are also compared. Finally, Section 4 summarizes the study and concludes this paper.

2. Materials and Methods

2.1. Data Description

A dataset used in BCI competition III (Dataset I) was used to investigate the effects of data augmentation on classification [23]. The subject imagined the movement of a left small finger (Class 1) and tongue (Class 2) for 3 s. The brainwave data (i.e., electrocorticography) were measured at a 1000 Hz sampling frequency. From the original dataset, 160 samples (80 samples for finger and tongue each) were used in the training dataset, and 100 samples (50 samples per class) were used to obtain the test accuracy. Among the 64 channels in the experiment, channel 30 was used in this study because it was located where brainwaves related to motor imagery were measured.

2.2. Methods

This study focused on the effects of random noise, used in CMD, on data augmentation. Thus, the GND for the random noise used in this study is introduced first. Subsequently, the CMD procedure is described.

2.2.1. Random Noise

The stochastic nature of random noise, which is generated by a normal distribution, is restricted because it only depends on its mean value and variance. To address this issue, GND [24,25] was used as the random noise in this study because the distribution of the original brainwave data was similar to that of GND. Specifically, the amplitudes of the frequency components are widely used to classify motor imagery brainwaves. Thus, the time series brainwave data were transferred to the frequency domain. Figure 1a,b show that the distributions of some frequencies had a blunt shape. This distribution was created using GNDv1. Additionally, the probability distributions (PDFs) of the amplitudes of some frequencies are similar to a skewed normal distribution, as shown in Figure 1c,d. This distribution was created using GNDv2. Considering these observations, GNDv1 and GNDv2 were chosen to produce random noise in this study.
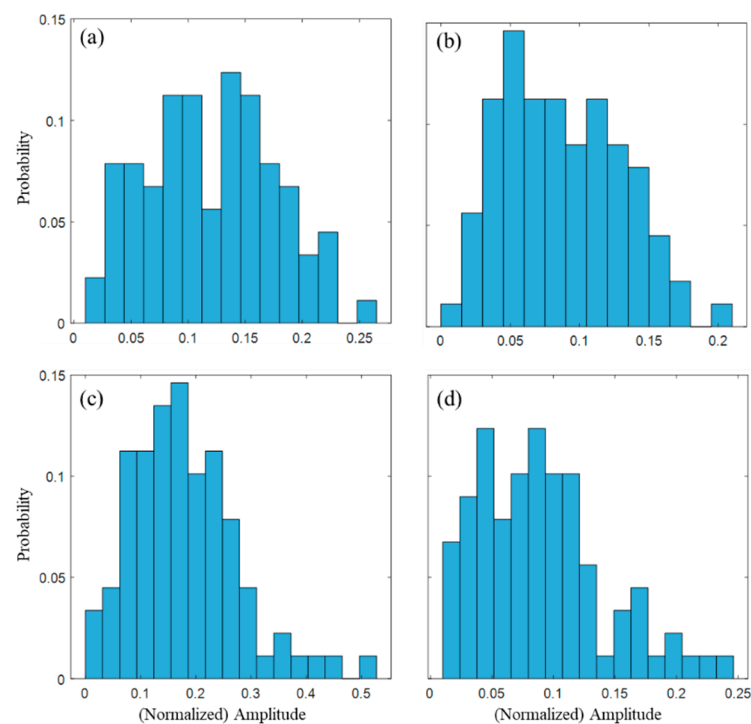
Figure 1. Distribution of amplitudes for (**a**) 24 Hz, (**b**) 29.7 Hz, (**c**) 19.3 Hz, and (**d**) 27.7 Hz. The amplitude was normalized by the signal power.

A random noise x following a specific PDF p_x can be generated using its cumulative distribution function (CDF) c_x [26] as follows. First, a random variable y is generated, which is equally distributed between 0 and 1. Then, x can be calculated as $x = c_x^{-1}(y)$. Thus, the inverse function of the CDF for the GND is required to generate random noise following a GND.

The PDF of GNDv1 is defined as:

$$p_{v1}(x) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left(-\left(\frac{|x-\mu|}{\alpha}\right)^{\beta}\right), \tag{1}$$

where β is a shape parameter that changes the kurtosis of the PDF; μ is the mean; σ is the standard deviation; α is a scale parameter defined as $\sqrt{\Gamma(\frac{1}{\beta})/\Gamma(\frac{3}{\beta})} \times |\sigma|$, where Γ is the gamma function. This distribution is the same as the normal distribution when $\beta = 2$, as shown in Figure 2a. When $\beta < 2$, p_{v1} is sharper than the PDF of the normal distribution, and vice versa.
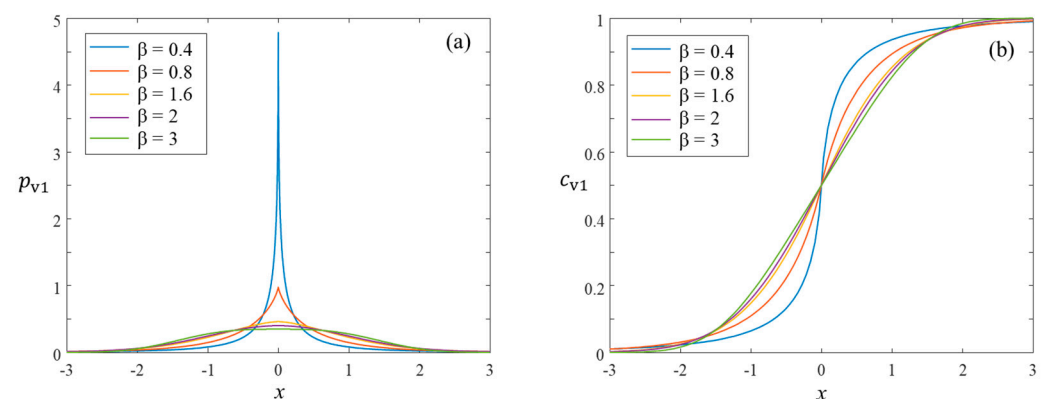

Figure 2. (**a**) Probability and (**b**) cumulative distributions for GNDv1 with zero mean and unit standard deviation.

A function c_{v1}, which is the CDF of GNDv1, can be obtained by integrating p_{v1} as follows:

$$c_{v1}(x) = \frac{1}{2} + \frac{\text{sgn}(x - \mu)}{2} \times g(|x - \mu|^{\beta}, \frac{1}{\beta}, \alpha^{\beta}), \tag{2}$$

where $g(h, a, b)$ is defined as $\frac{1}{b^a \Gamma(a)} \int_0^h s^{a-1} \exp\left(-\frac{s}{b}\right) ds$ and sgn denotes the sign function. Then, the inverse function of c_{v1} can be calculated as:

$$c_{v1}^{-1}(y) = \mu + \text{sgn}(c_{v1}(y) - 0.5)\left\{g^{-1}\left(2|c_{v1}(y) - 0.5|, \frac{1}{\beta}, \alpha^{\beta}\right)\right\}^{\frac{1}{\beta}}. \tag{3}$$

This inverse function is used to generate random noise following GNDv1.

The random noise following GNDv2 is determined using the shape parameter κ. Its PDF is skewed to the right when κ is positive, and vice versa. Specifically, the PDF of GNDv2 can be represented as:

$$p_{v2}(x) = \frac{1}{\sigma \sqrt{2\pi}\{\alpha - \kappa(x - \xi)\}} \exp\left(-\frac{1}{2}\left(\frac{z(x) - \mu}{\sigma}\right)^2\right), \tag{4}$$

where z is defined as $-\frac{1}{\kappa} \log\left(1 - \frac{\kappa(x - \xi)}{\alpha}\right)$ when $\kappa \neq 0$ and as $\frac{x - \xi}{\alpha}$ when $\kappa = 0$.

Then, the CDF of GNDv2 can be calculated as:

$$c_{v2}(x) = \frac{1}{2}[1 + \text{erf}(\frac{z(x) - \mu}{\sigma \sqrt{2}})], \tag{5}$$

where $\text{erf}(x)$ is the error function defined as $\frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

As κ increases, the PDF peak shifts to the right, as shown in Figure 3. The inverse function of c_{v2} can be calculated as:

$$c_{v2}^{-1}(y) = \xi + \frac{\alpha}{\kappa}\{1 - \exp(-\kappa h(y) + \mu)\}, \tag{6}$$

where α and ξ are defined as $\frac{|\sigma\kappa|}{\sqrt{e^{2\kappa^2} - e^{\kappa^2}}}$ and $\frac{\alpha}{\kappa} e^{\kappa^2}(e^{\kappa^2} - 1)$, respectively, and $h(y)$ is determined as $\sigma \sqrt{2} \text{erf}^{-1}(2y - 1)$. This inverse function is used to generate random noise for the CMD.
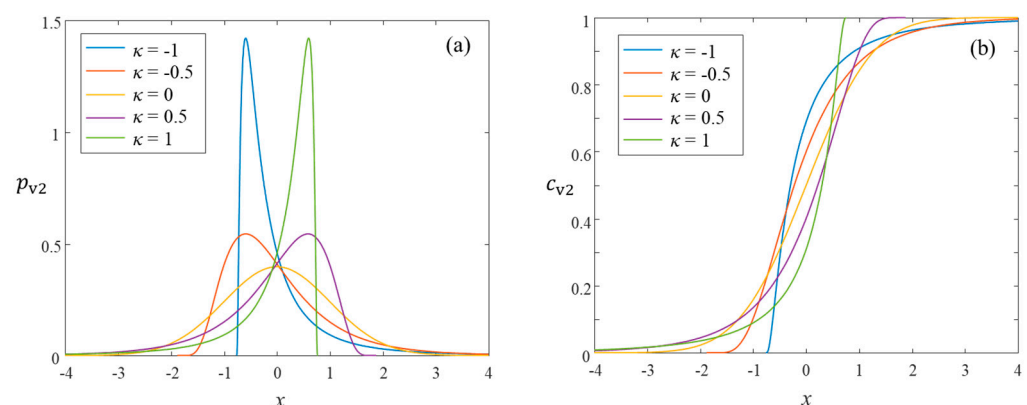

Figure 3. (**a**) Probability and (**b**) cumulative distributions for GNDv2 with zero mean and unit standard deviation.

2.2.2. Covariance Matrix Decomposition

An artificial dataset, with a distribution similar to that of the original dataset, was created using CMD. Let the original and artificial datasets be $D \in R^{k \times N}$ and $D_c \in R^{k \times n}$, respectively, where N is the number of original data points; n is the number of augmented

data points; k is the number of features. The ith sample of the original dataset can be expressed as $d^{(i)} = u + v^{(i)}$, and the ith sample of the artificial dataset can be expressed as $d_a^{(i)} = u + v_a^{(i)}$, where u is a vector composed of mean values of Y, and $v^{(i)}$ and $v_c^{(i)}$ are the stochastic components. Next, suppose that X is a matrix composed of random noise vectors as $X = \left[x^{(1)} \, x^{(2)} \ldots x^{(n)} \right]$, where $x^{(i)}$ is a random noise vector with zero mean and unit standard deviation. $\mathrm{Cov}[\cdot]$ can be defined as an operator that calculates the covariance matrix of the input vectors. If there exists a matrix L, which relates $v_c^{(i)}$ and random noise as $v_c^{(i)} = Lx^{(i)}$, the covariance matrix of Y_c is equal to LL^T (Proof: $\mathrm{Cov}\left[LX \cdot (LX)^T \right] = \mathrm{Cov}\left[LXX^T L^T \right] = L\left(\mathrm{Cov}\left[XX^T \right] \right)L^T = LIL^T = LL^T$, where I is the identity matrix). It is worth noting that this relation can be derived because the mean and standard deviation of the random noise are 0 and 1, respectively. To generate artificial data similar to the original data, the covariance matrices of Y and Y_c should be the same. Thus, L can be obtained when LL^T is the same as the covariance matrix of the original dataset.

Figure 4 briefly presents the data augmentation pipeline used in this study. The artificial data generated by CMD and GND were combined with the original data in the training stage. Then, both were used for training. In the test stage, only the original data for testing were used to obtain the classification accuracy.
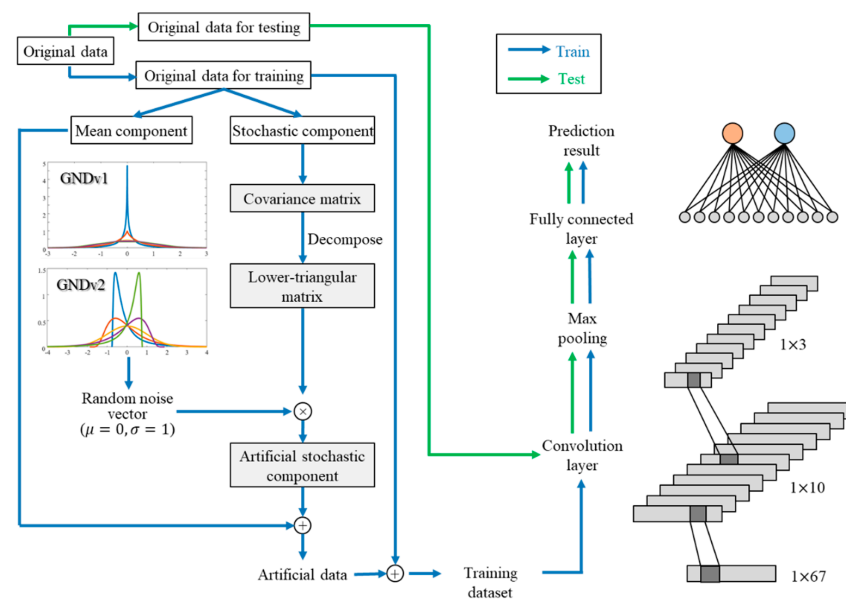

Figure 4. Training and test pipeline using data augmentation via CMD.

Frequency domain features are widely used to classify motor imagery brainwaves; therefore, the following pre-processing was applied in this study. Time series samples were transferred to the frequency domain via a fast Fourier transform. Then, amplitudes with frequencies between 8 and 16 Hz were collected, as this frequency range is closely related to motor imagery brainwaves [27]. Next, the amplitude values were normalized by integrating the amplitudes over the frequency.

2.2.3. Convolutional Neural Network

The convolutional neural network (CNN) structure used in a previous study [21] was adopted. Specifically, a single CNN layer composed of 16 (1×10) convolution kernels and 1×3 max pooling layers was used. The LeakyReLU activation function was applied after the convolution operations, and softmax was applied to the output nodes. The Adam optimizer was used for training, and the learning rate, decay rate for the first moment, and decay rate for the second moment were set to 0.0002, 0.9, and 0.999, respectively. Training was conducted for 10,000 epochs.

Normalized amplitudes of frequencies between 8 and 16 Hz were used as the input to the CNN. Since the trained model acted as a binary classifier, it could distinguish whether the input test data signal belonged to Class 1 or Class 2.

3. Results and Discussion

3.1. Data Augmentation

The augmented data were affected by the PDF of the random noise, as shown in Figures 5 and 6. The GNDv1 with a small β exhibited a high central density, as shown in Figure 1a. Subsequently, most artificial data samples tended to be close to each other, as shown in Figure 5b. Furthermore, the GNDv1 with a small β exhibited a higher probability density at its tail than that with a large β. Thus, few samples exhibited extreme fluctuations, as shown in Figure 5b. When β was large, the samples were widely distributed, as shown in Figure 5c,d. However, extremely fluctuating samples were not generated for a large β.

The GNDv2 with a negative κ had a higher density on the negative side than on the positive side. Thus, the generated samples tended to be skewed toward lower values, as shown in Figure 6b. Additionally, a negative κ yielded a higher density, at larger positive values, than a positive κ, resulting in occasional fluctuation to the positive side in a few samples. Conversely, when κ was positive, most generated samples were skewed toward higher values, as shown in Figure 6d; this was because the PDF of the random noise was skewed toward the positive side. Some artificial samples of positive κ exhibited negative amplitudes because the corresponding GNDv2 had a considerable probability density for negative random noise, as shown in Figure 3a. Thus, negative amplitudes could downgrade the classification accuracy because they were not observed in the original data.
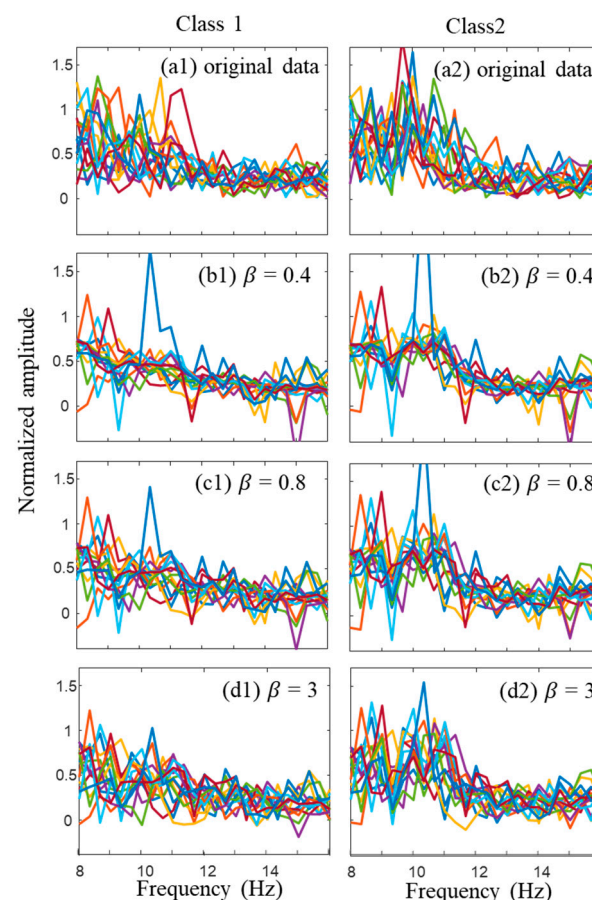

Figure 5. Artificial data generated with GNDv1. The *y*-axis represents the normalized amplitude in the frequency domain.
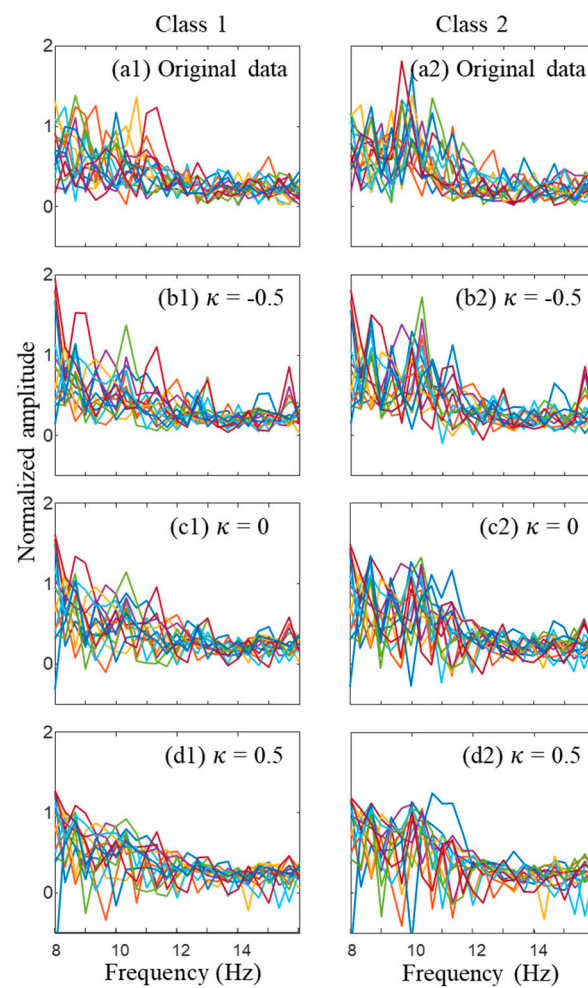
Figure 6. Artificial data generated with GNDvw2. The *y*-axis represents the normalized amplitude in the frequency domain.

A cut-off on the PDF of the random noise can be used to prevent negative amplitudes in the artificial data. Specifically, random noise values less than a threshold e_{th} can be excluded when augmenting the data, as shown in Figure 7. Figure 8 shows the augmented data when $\kappa = 0.5$ and $e_{th} = -2, -1$, and -0.5. When $e_{th} = -0.5$ and -1, no negative amplitudes were generated. Moreover, they exhibited a pattern similar to that of the original data.
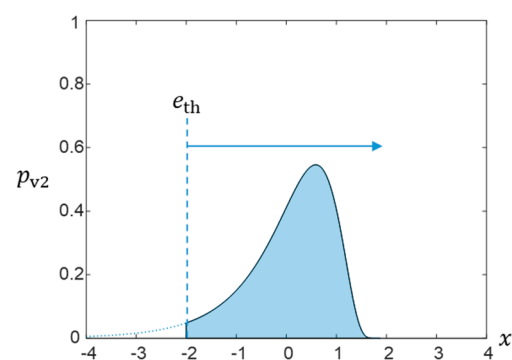

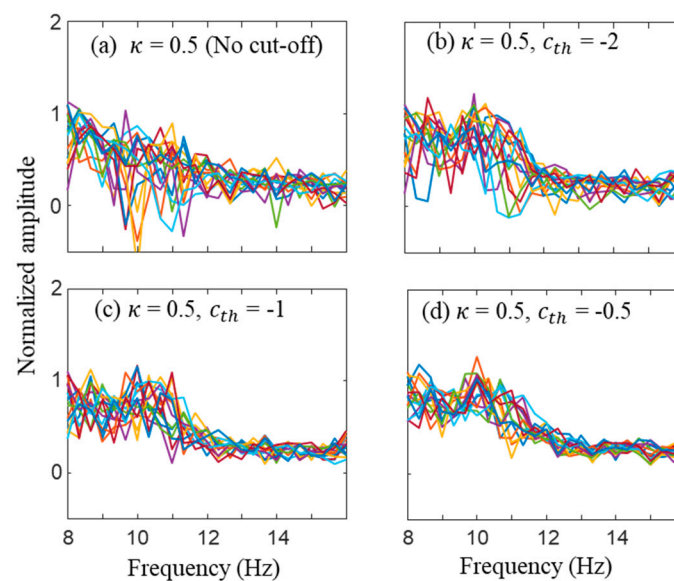**Figure 7.** Cut-off for the PDF of GNDv2.

Figure 8. Augmented data of Class 1 with GNDv2: (**a**) $\kappa = 0.5$ and no cut-off; (**b**) $\kappa = 0.5$ and $e_{th} = -2$; (**c**) $\kappa = 0.5$ and $c_{th} = -1$; (**d**) $\kappa = 0.5$ and $e_{th} = -0.5$.

To investigate the effects of this cut-off technique on classification accuracy, a CNN classifier was trained with the augmented data. Then, the accuracy values in the absence/presence of the cut-off were obtained and compared, as presented in Table 1. When the cut-off was applied, the augmented data presented greater similarity to the original data; therefore, we expected the accuracy to increase over e_{th}. However, the accuracy decreased as more of the PDF was removed.

Table 1. Effects of data cut-off on classification accuracy. The GNDv2 with $\kappa = 0.5$ was used. The numbers in parentheses represent the standard errors.

Number of Augmented Data Per Class	e_{th}			
	No Cut-Off	-2	-1	-0.5
0 (Original)			71.2 (0.663)	
80	74.12 (0.543)	72.48 (0.473)	71.2 (0.476)	70.04 (0.531)
160	76.52 (0.444)	74.48 (0.724)	71.84 (0.594)	70.80 (0.493)
480	79.72 (0.604)	75.08 (0.557)	71.68 (0.340)	70.40 (0.400)
800	78.08 (0.476)	76.68 (0.403)	71.64 (0.500)	69.00 (0.370)

This result was obtained because sample variation was important for accurate classification. The sample variation decreased significantly when the cut-off was applied to the random noise, as shown in Figure 8. Additionally, the number of low-quality samples (i.e., samples with negative amplitudes) is very small; thus, their effect on the classification can be deemed negligible. Consequently, the cut-off technique was not used.

3.2. Accuracy Improvement

The accuracy was obtained for various shape parameters over different augmented data samples; the shape parameter β of GNDv1 was varied from 0.2 to 3.5, while the parameter κ of GNDv2 was varied from -2 to 2. A total of 80, 160, 480, and 800 artificial samples were generated per class. Then, the artificial and original datasets were used for training the CNN classifier.

For a reliable comparison, the effects of the initial weight of the CNN classifier and the augmentation randomness on accuracy need to be minimized. To this end, the accuracies for various initial weights and several augmentation trials were obtained. The final accuracy

value was determined by averaging these accuracies. For example, to obtain the averaged accuracy for artificial samples with $\beta = 0.2$ and 80, five datasets with 80 artificial samples were created with $\beta = 0.2$. Then, five CNN classifiers (with the same structure and different initial weights) were trained using the first artificial dataset and the original dataset. Next, the five CNN classifiers were trained using the second artificial dataset and the original dataset. As this training process was conducted using all artificial datasets, a total of 25 accuracy values were obtained. Then, the average and standard error of the accuracy for artificial samples with $\beta = 0.2$ and 80 can be calculated. Figures 9 and 10 show the average and standard error of the accuracy obtained via the described method for GNDv1 and GNDv2, respectively.
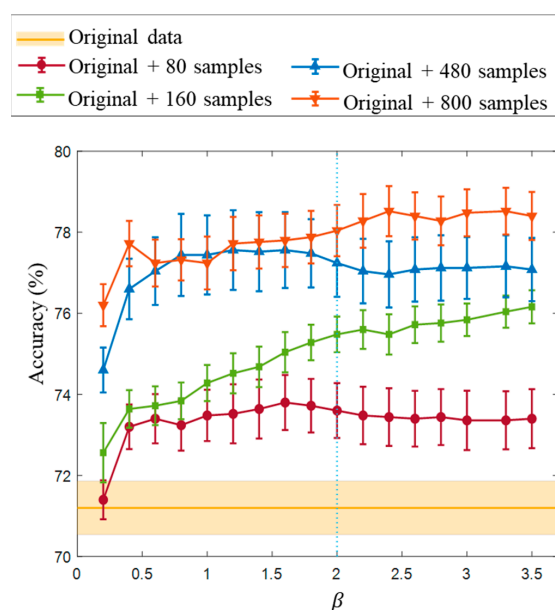
Figure 9. Classification accuracy when GNDv1 is used; the symbols and error bars represent the average accuracy and standard error, respectively.
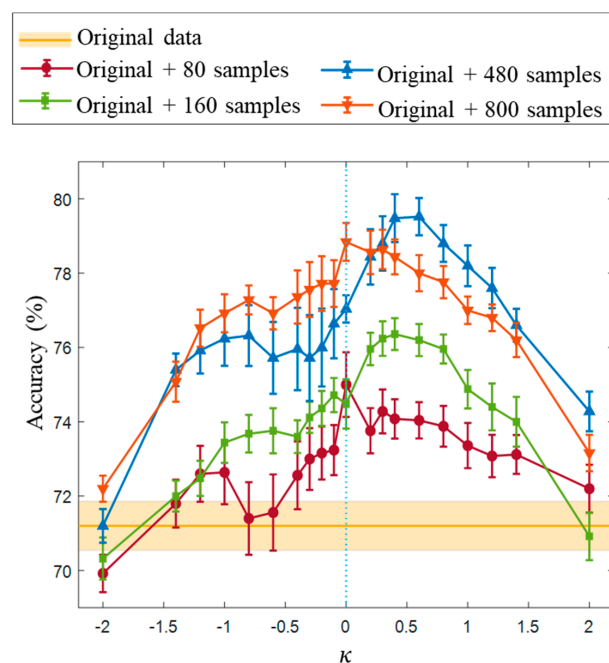
Figure 10. Classification accuracy when GNDv2 is used; the symbols and error bars represent the average accuracy and standard error, respectively.

When the augmented dataset was added to the training dataset, the accuracy significantly increased in most cases compared to when the original dataset was used alone. In the case of the dataset generated using GNDv1 (Figure 9), the accuracy increases as β increases for $\beta < 1$. This suggests that a smoother distribution is more effective for generating high-quality brainwave data. For $\beta > 1$, the effects of β on accuracy were negligible because the changes in the PDF were not considerable, as shown in Figure 2a. Furthermore, the accuracy increased as the number of samples with the artificial dataset increased. For example, for a large β, the accuracy improvement was approximately 2% when 80 samples were augmented. The improvement was approximately 7% when 800 samples were added.

When artificial data were created using GNDv2, high accuracy was achieved for $0 < \kappa < 1$, as shown in Figure 10. This suggests that high-quality brainwave signals can be generated when the PDF of the random noise is slightly skewed toward the positive side. The accuracy also improved as the number of artificial samples increased. However, when $\kappa > 0.3$, a higher level of accuracy was achieved with 480 samples than with 800 samples. We speculate that this counterintuitive result was obtained because the effect of negative amplitudes was considerable only when the number of artificial samples was very large. Specifically, when κ was positive, negative amplitudes were generated, which were not observed in the original dataset. However, negative amplitudes would be rare if the number of samples was not sufficiently large (i.e., less than 480). Thus, the effect of negative amplitudes was small. Therefore, if a sufficiently large number of samples is created, the training dataset is likely to contain a considerable number of negative amplitudes, leading to accuracy degradation. Thus, a lower accuracy was achieved with 800 artificial samples than with 480 samples.

To verify that the accuracy could be improved by the flexible CMD in other biosignals, the model was applied to an electromyography (EMG) signal. Specifically, EMG signals corresponding to three holding motions, namely holding a glass of water, water bottle, and pen, in the public EMG dataset [28] were used in this study. The data preprocessing and data augmentation techniques used for brainwave signals were also applied to the EMG signals. Figures 11 and 12 show the accuracies obtained via GND v1 and GNDv2, respectively.
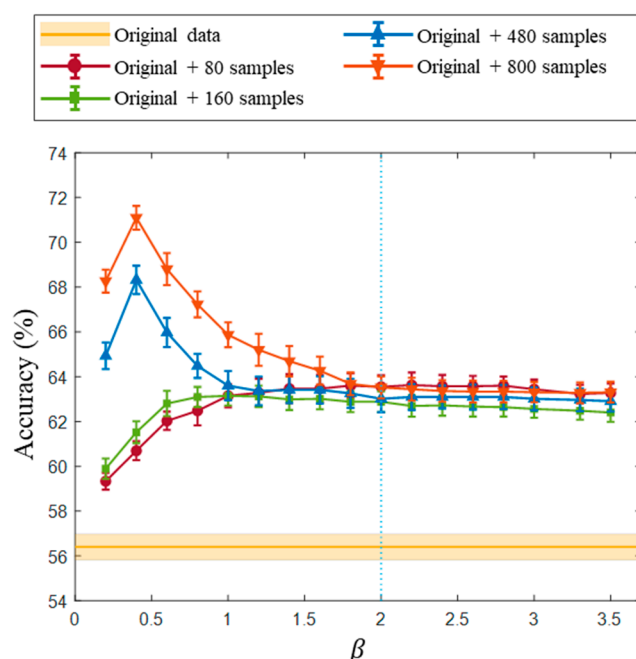

Figure 11. Classification accuracy of EMG data with GNDv1; the symbols and error bars represent the average accuracy and standard error, respectively.
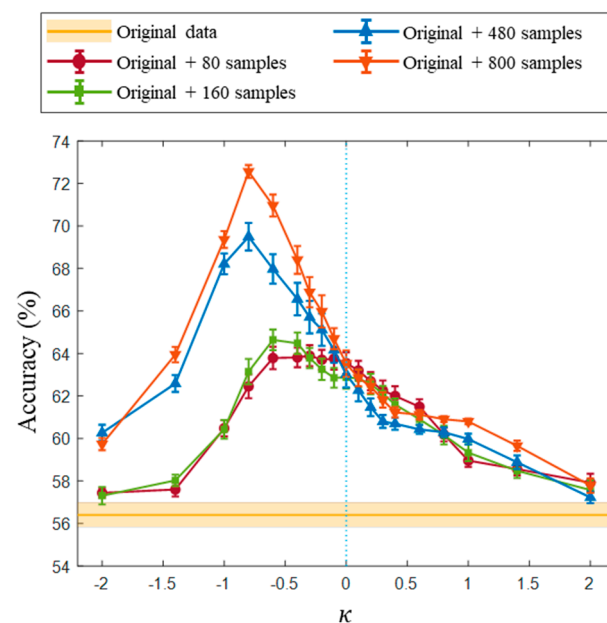
Figure 12. Classification accuracy of EMG data with GNDv2; the symbols and error bars represent the average accuracy and standard error, respectively.

The augmented dataset significantly increased the accuracy of EMG signals. When 800 samples were used, the maximum accuracy was approximately 71.1%, which was very high considering that the accuracy for a normal distribution was approximately 63.5%. When 80 and 160 samples were augmented with GNDv1, the accuracy increased over β and saturated when β was approximately 1. Meanwhile, when 480 or 800 samples were augmented with GNDv1, the maximum accuracy was obtained when β was 0.4. This difference can be attributed to an insufficient number of random noises. When the number of generated random noises is small, the actual distribution of the generated random noise can be different from the expected distribution (i.e., GNDv1). Thus, the effects of β on the accuracy are different for small and large samples.

When artificial EMG data were created using GNDv2, a high accuracy was also obtained, as shown in Figure 12. When 800 samples were used, the maximum accuracy was approximately 72.6%, which was significantly higher than the value obtained with the normal distribution. Moreover, the accuracy tends to be high for negative κ, suggesting that high-quality EMG signals can be artificially generated when random noise distribution is slightly skewed to the left. The value of κ, corresponding to the maximum accuracy, is different for small (i.e., 80 and 160) and large (i.e., 480 and 800) samples. This difference can also be explained by the actual distribution of generated random noise.

4. Conclusions

In previous studies, normally distributed random noise was used for CMD. Therefore, in this study, an artificial dataset was created using other types of random noise distributions. To predict the type of motor imagery brainwave more accurately, a CNN classifier was trained with the dataset augmented via CMD based on GNDv1 and GNDv2. The PDF parameters of the random noise were varied, and their effects were investigated. The accuracy improved for larger β in GNDv1 and for small positive κ in GNDv2. In addition, the classification accuracy tended to increase when more artificial data were generated.

It is worth noting that the optimal parameter values for β and κ can be changed if the proposed method is applied to other bio-signals. Thus, to determine the optimal parameters for random noise, the accuracy of the validation dataset needs to be obtained and compared with various parameter values.

This study has proposed a generic method to release the constraints of classical CMD using GND. This method can be used to generate other types of bio-signals. However, GND might not be an optimal PDF for random noise for certain target signals. Other types of PDFs, such as multimodal distributions or Student's t-distributions, can be used for data augmentation in future work. Furthermore, a more general PDF for random noise, which is defined by many parameters, can provide higher-quality artificial data, the parameters of which can be trained to generate artificial data such that their distribution is similar to that of the original data.

Author Contributions: Conceptualization, W.N.; methodology, H.L.; software, H.L.; validation, H.L.; formal analysis, H.L.; investigation, H.L.; resources, W.Y.; data curation, H.L.; writing—original draft preparation, H.L.; writing—review and editing, W.N. and W.Y.; visualization, H.L.; supervision, W.N.; project administration, W.N.; funding acquisition, W.N. All authors have read and agreed to the published version of the manuscript.

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

References

1. Shih, J.J.; Krusienski, D.J.; Wolpaw, J. *Brain-Computer Interfaces in Medicine*; Elsevier: Amsterdam, The Netherlands, 2012; Volume 87, pp. 268–279.
2. Wolpaw, J.R. Brain—Computer interfaces. In *Handbook of Clinical Neurology*; Elsevier: Amsterdam, The Netherlands, 2013; pp. 67–74.
3. Abdulkader, S.N.; Atia, A.; Mostafa, M.-S. Brain computer interfacing: Applications and challenges. *Egypt. Inform. J.* **2015**, *16*, 213–230. [CrossRef]
4. Daly, J.J.; Wolpaw, J. Brain—computer interfaces in neurological rehabilitation. *Lancet Neurol.* **2008**, *7*, 1032–1043. [CrossRef]
5. Bonnet, L.; Lotte, F.; Lecuyer, A. Two Brains, One Game: Design and Evaluation of a Multiuser BCI Video Game Based on Motor Imagery. *IEEE Trans. Comput. Intell. AI Games* **2013**, *5*, 185–198. [CrossRef]
6. Zhang, R.; Li, Y.; Yan, Y.; Zhang, H.; Wu, S.; Yu, T.; Gu, Z. Control of a Wheelchair in an Indoor Environment Based on a Brain–Computer Interface and Automated Navigation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2015**, *24*, 128–139. [CrossRef] [PubMed]
7. Hazrati, M.K.; Erfanian, A. An online EEG-based brain–computer interface for controlling hand grasp using an adaptive probabilistic neural network. *Med. Eng. Phys.* **2010**, *32*, 730–739. [CrossRef] [PubMed]
8. Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J.; Pfurtscheller, G.; Vaughan, T.M. Brain–computer interfaces for communication and control. *Clin. Neurophysiol.* **2002**, *113*, 767–791. [CrossRef]
9. Yin, Z.; Zhang, J. Cross-subject recognition of operator functional states via EEG and switching deep belief networks with adaptive weights. *Neurocomputing* **2017**, *260*, 349–366. [CrossRef]
10. Wang, F.; Zhong, S.-H.; Peng, J.; Jiang, J.; Liu, Y. *Data Augmentation for EEG-Based Emotion Recognition with Deep Convolutional Neural Networks*; Lecture Notes in Computer Science; 2018; pp. 82–93.
11. Salama, E.S.; El-Khoribi, R.A.; Shoman, M.E.; Wahby, M.A. EEG-Based Emotion Recognition using 3D Convolutional Neural Networks. *Int. J. Adv. Comput. Sci. Appl.* **2018**, *9*. [CrossRef]
12. Parvan, M.; Ghiasi, A.R.; Rezaii, T.Y.; Farzamnia, A. Transfer Learning based Motor Imagery Classification using Convolutional Neural Networks. In Proceedings of the 27th Iranian Conference on Electrical Engineering (ICEE), Yazd, Iran, 30 April–2 May 2019; pp. 1825–1828.
13. Piplani, T.; Merill, N.; Chuang, J. Faking it, Making it: Fooling and Improving Brain-Based Authentication with Generative Adversarial Networks. In Proceedings of the IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), Redondo Beach, CA, USA, 22–25 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–7.
14. Hartmann, K.G.; Schirrmeister, R.T.; Ball, T. EEG-GAN: Generative adversarial networks for electroencephalograhic (EEG) brain signals. *arXiv* **2018**, arXiv:1806.01875. Available online: https://arxiv.org/abs/1806.01875 (accessed on 7 October 2021).
15. Luo, Y.; Lu, B.-L. EEG Data Augmentation for Emotion Recognition Using a Conditional Wasserstein GAN. In Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; IEEE: Piscataway, NJ, USA, 2018; Volume 2018, pp. 2535–2538.
16. Zhang, Q.; Liu, Y. Improving brain computer interface performance by data augmentation with conditional deep convo-lutional generative adversarial networks. *arXiv* **2018**, arXiv:1806.07108.
17. Ni, X. Research of data mining based on neural networks. *World Acad. Sci. Eng. Technol.* **2008**, *39*, 381–384.
18. Rue, H.; Held, L. *Gaussian Markov Random Fields: Theory and Applications*; CRC Press: Boca Raton, FL, USA, 2005.

19. Chenari, R.J.; Farahbakhsh, H.K. Generating non-stationary random fields of auto-correlated, normally distributed CPT profile by matrix decomposition method. *Georisk: Assess. Manag. Risk Eng. Syst. Geohazard.* **2015**, *9*, 96–108. [CrossRef]

20. Okaze, T.; Mochida, A. Cholesky decomposition–based generation of artificial inflow turbulence including scalar fluctuation. *Comput. Fluids* **2017**, *159*, 23–32. [CrossRef]

21. Yang, W.; Nam, W. Brainwave Classification Using Covariance-Based Data Augmentation. *IEEE Access* **2020**, *8*, 211714–211722. [CrossRef]

22. Yang, W.; Nam, W. Data synthesis method preserving correlation of features. *Pattern Recognit.* **2021**, in press. [CrossRef]

23. BCI Competition III Dataset. Available online: http://www.bbci.de/competition/iii/desc_I.html (accessed on 7 October 2021).

24. Nadarajah, S.A. generalized normal distribution. *J. Appl. Stat.* **2005**, *32*, 685–694. [CrossRef]

25. Hosking, J.R.M.; Wallis, J.R. *Regional Frequency Analysis: An Approach Based on L-Moments*; Cambridge University Press: Cambridge, UK, 2005.

26. Steele, J.M. Non-Uniform Random Variate Generation (Luc Devroye). *SIAM Rev.* **1987**, *29*, 675–676. [CrossRef]

27. Jeon, Y.; Nam, C.S.; Kim, Y.-J.; Whang, M.C. Event-related (De)synchronization (ERD/ERS) during motor imagery tasks: Implications for brain–computer interfaces. *Int. J. Ind. Ergon.* **2011**, *41*, 428–436. [CrossRef]

28. Sapsanis, C. Recognition of Basic Hand Movements Using Electromyography. Master's Thesis, Johns Hopkins University, Baltimore, MD, USA, June 2013.