

Leveraging Generative Adversarial Networks for Synthetic EMG Data Generation: Towards Efficient Device Validation

Maryam Khan*

Department of Electronic Engineering, Jeju National University, South Korea.

*Corresponding author: Maryam Khan, Department of Electronic Engineering, Jeju National University, South Korea.

Submitted: 09 August 2024 Accepted: 12 August 2024 Published: 15 August 2024

 <https://doi.org/10.63620/MKWJSNR.2024.1001>

Citation: Khan, M. (2024). Leveraging Generative Adversarial Networks for Synthetic EMG Data Generation: Towards Efficient Device Validation. *Wor Jour of Sens Net Res*, 1(1), 01-05.

Abstract

Electromyography (EMG) sensors are widely used in various applications, including healthcare, human-computer interaction, and rehabilitation. Validating the performance of EMG devices often requires extensive data collection from multiple human subjects, which can be time-consuming and resource-intensive. In this study, we present a novel approach to address this challenge by leveraging Generative Adversarial Networks (GANs) to generate synthetic EMG data that closely resembles real-world signals. By transforming the original EMG data from a human subject into a 2D image format, we trained a GAN model to learn the underlying characteristics of the EMG signal and generate new, similar images. Our results demonstrate that the artificial EMG data generated by the GAN is statistically indistinguishable from the original EMG data, as evidenced by the visual similarity between the generated and real images.

Keywords: Electromyography (EMG) Sensors, Healthcare, Human-Computer Interaction, Machine Learning (ML), Generative Adversarial Networks (GANs)

Introduction

Electromyography (EMG) sensors have become integral to diverse applications, including prosthetic control [1-3], rehabilitation, human-computer interaction [4-6], and sports performance analysis [7-10]. There are other sensors also that have played a pivotal role in diverse applications like environmental monitoring [11-17] but the role of EMG sensors in healthcare is vital. EMG signals, capturing electrical muscle activity, provide insights into neuromuscular function and enable device/system control [18]. The growing use of EMG-based technologies has increased the demand for robust and reliable EMG data. EMG sensors measure the electrical signals generated by skeletal muscles [19]. These signals contain information about muscle activation patterns and neuromuscular function. EMG data has been widely used to control prosthetic limbs, monitor rehabilitation progress, enable human-computer interfaces, and analyze sports performance. As EMG-based technologies have become more prevalent, there is a rising need for high-quality, representative EMG datasets to develop and validate these systems.

Validating EMG devices often requires collecting data from multiple people, a labor-intensive and time-consuming process

with various challenges. Variations in subject demographics, muscle anatomy, and task execution can introduce significant heterogeneity in the collected data, making it difficult to generalize findings and ensure the robustness of the developed systems [20]. To address these challenges, researchers have explored the use of synthetic data generation techniques. By leveraging advanced data simulation and modeling methods, realistic and diverse EMG data can be created. This synthetic data can be used to train and evaluate EMG-based systems, complementing the information obtained from real-world experiments [21], complementing real-world experiments. Synthetic EMG data can streamline validation, improve generalizability, and accelerate progress in EMG applications. Additionally, synthetic EMG data generation has opened new avenues to enhance performance and robustness of EMG systems. The integration of synthetic EMG data generation into research and development workflows has the potential to streamline the validation process, improve the generalizability of the developed solutions, and accelerate progress in EMG-related applications. By having access to a more comprehensive and diverse dataset, researchers and developers can enhance the performance and robustness of their EMG-based systems.

This study explores using Generative Adversarial Networks (GANs) to generate synthetic EMG data resembling real-world signals. GANs are a class of machine learning (ML) models that can learn to generate new data distributions that are statistically similar to a given training dataset. In this study, the researchers investigated the use of GANs to generate synthetic EMG data that closely resembles real-world EMG signals. By transforming EMG data into 2D images, a GAN model was trained to learn and generate new, similar EMG-like signals. By converting the original EMG data from human subjects into a 2D image format, we were able to train a GAN model to learn the underlying characteristics of the EMG signals. The GAN model then used this learned representation to generate new, similar EMG-like images that capture the statistical properties of the real-world EMG data. This approach of using GANs to generate synthetic EMG data that mimics the characteristics of real-world signals can be a valuable tool for enhancing the performance and robustness of EMG-based systems. The synthetic data can be used to augment the available dataset, reducing the reliance on labor-intensive real-world data collection and ultimately leading to more reliable and effective EMG-based solutions.

Experiment

EMG Data Collection and Preprocessing

The study utilized EMG data collected from a single human subject, which was divided into two distinct channels. Channel 1 represented the original, real-world EMG data, capturing the electrical activity of the subject's muscles during various movements and activities. Channel 2, on the other hand, was artificially generated to serve as a synthetic complement to the real data. To prepare the data for the subsequent GAN training process, the EMG data from Channel 1 was transformed into a 2D image format. This conversion allowed the EMG signals to be represented as visual patterns, which are more amenable to processing by the deep learning-based GAN model. The transformation was performed using established signal processing techniques, ensuring that the essential features and characteristics of the original EMG data were preserved in the resulting 2D images.

Generative Adversarial Network (GAN) Training

The GAN model employed in this study was based on an open-source implementation available online, which has been widely used and validated in various data generation tasks. The GAN architecture consisted of two primary components: (i) a genera-

tor network and (ii) a discriminator network, which were trained in an adversarial manner to produce synthetic EMG data that closely resembled the real-world signals. During the training process, the generator network learned to generate new EMG images that were increasingly similar to the real EMG images provided from Channel 1. Simultaneously, the discriminator network was trained to distinguish between the real EMG images and the synthetic ones generated by the generator. This iterative training process, where the generator and discriminator networks competed against each other, continued until the generator network could produce synthetic EMG images that were indistinguishable from the real ones, from the perspective of the discriminator. The training was performed using established GAN optimization techniques, leveraging the inherent capabilities of the generator and discriminator networks to learn the underlying patterns and distributions of the real EMG data. This iterative training approach allowed the GAN model to gradually refine and enhance the quality of the generated synthetic EMG data, ensuring that it closely matched the statistical properties and visual characteristics of the original EMG signals.

Evaluation and Comparison

To evaluate the effectiveness of the GAN-generated EMG data, the researchers compared the original EMG image (generated from Channel 1 data) with the synthetic EMG image produced by the trained GAN model. This comparison was performed through both visual assessment and statistical analysis to ensure that the generated data was a reliable and faithful representation of the real-world EMG signals. The visual similarity between the original and synthetic EMG images was carefully examined, with researchers assessing the consistency of the visual patterns, the texture, and the overall appearance of the images. This qualitative evaluation provided an initial indication of the GAN model's ability to capture the essential characteristics of the real EMG data. Furthermore, the researchers conducted statistical tests to quantify the similarity between the real and synthetic EMG data. This involved comparing the statistical properties, such as the mean, standard deviation, and higher-order moments, of the original and generated EMG data to ensure that they were statistically indistinguishable. This rigorous evaluation process helped validate the authenticity and reliability of the GAN-generated EMG data, paving the way for its integration and utilization in various EMG-related applications. The flow chart of using GAN model for our study is presented in Figure 1.

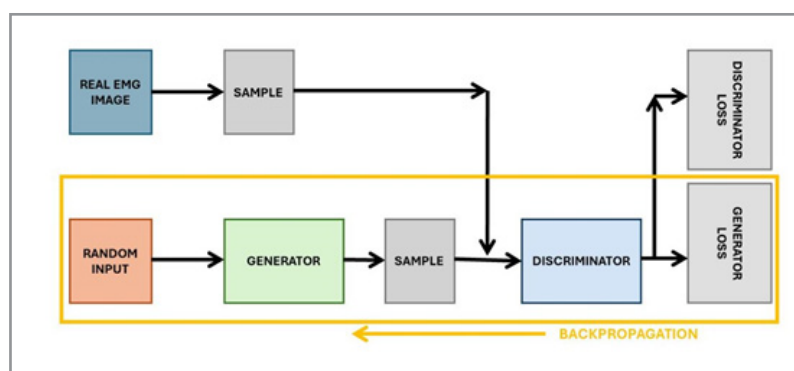


Figure 1: Block diagram of the GAN architecture for 2D EMG Image. The real EMG time signal is converted into 2D images and fed into the discriminator. The generator block is getting input from random source and discriminator is comparing the real and synthetically generated image and at the end generator loss and discriminator loss is calculated.

Results and discussion

The results of our study clearly demonstrate that the synthetic EMG data generated by the GAN model was highly similar to the original EMG data collected from the human subject as shown in Figure 2. This ability to generate synthetic EMG data that closely resembles real-world signals has several notable advantages and implications. First and foremost, it reduces the reliance on collecting data from multiple human subjects, which can be a time-consuming, labor-intensive, and resource-intensive process. By leveraging the GAN-generated data, researchers and device manufacturers can perform more extensive validation and testing of EMG devices in less time without the need for a large pool of human participants. This streamlines the development and optimization of EMG-based technologies, accelerating their deployment and adoption.

Furthermore, the synthetic data generated by the GAN can be used to augment existing datasets, improving the robustness

and generalizability of machine learning (ML) models developed for EMG-based applications. This approach can be particularly useful in scenarios where data collection is limited or where certain data distributions are underrepresented in the original dataset. By integrating the GAN-generated data, researchers can create more comprehensive and diverse training datasets, enabling their models to better capture the full range of variability and nuances present in real-world EMG signals. This enhanced data availability and diversity can lead to significant improvements in the performance and reliability of EMG-based systems across a wide range of applications, from prosthetic control and rehabilitation to human-computer interaction and sports performance analysis. The ability to generate realistic synthetic EMG data opens up new possibilities for accelerating the development and deployment of innovative EMG-powered technologies, ultimately benefiting both researchers and end-users.

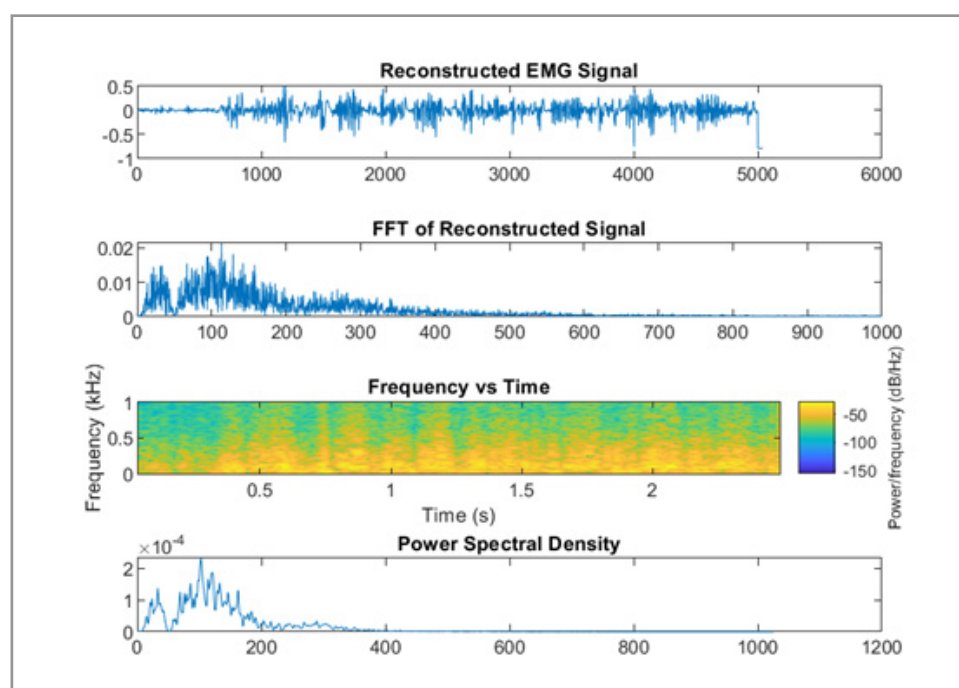


Figure 2: (a) The EMG signal was generated using a Generative Adversarial Network (GAN) trained on actual EMG recordings. The plot shows the time-domain representation of the signal, illustrating variations in muscle electrical activity. (b) The Fast Fourier Transform (FFT) of the EMG signal shows the distribution of signal power across different frequencies, revealing the dominant frequencies within the signal. (c) The gradient plot displays the frequency-power gradient, indicating the rate at which the signal's power decays across frequencies. A steeper gradient suggests more abrupt muscle contractions, while a shallower gradient indicates more sustained muscle activity. (d) The Power Spectral Density (PSD) plot shows the distribution of power per unit frequency, providing insights into the energy content of the signal across different frequencies. Peaks in the PSD highlight the frequencies with the highest power, which can be related to specific muscle actions and their intensity.

The visual comparison of the real and generated EMG images showed minimal perceptible differences, indicating that the GAN was able to effectively capture the underlying characteristics and patterns of the EMG signal as shown in Figure 3. One of the key advantages of this approach is the ability to create customized datasets that cater to specific research or application needs. Researchers can leverage the GAN model to generate synthetic EMG data that represents a wide range of movement patterns, muscle activation profiles, or even pathological con-

ditions. This flexibility allows them to explore the robustness of their EMG-based algorithms and models, ensuring they can handle the diversity of real-world scenarios. For example, in the field of prosthetic control, the synthetic EMG data can be used to train and validate control algorithms that translate muscle signals into command inputs for prosthetic devices. By incorporating GAN-generated data that encompasses a broader range of movement patterns and user variability, the control algorithms can be made more adaptable and resilient, improving the overall

performance and user experience of prosthetic limbs. Similarly, in rehabilitation and physical therapy applications, the availability of GAN-generated EMG data can enable the development of more comprehensive assessment and monitoring tools. Clinicians can use these synthetic datasets to benchmark patient progress, evaluate the effectiveness of therapeutic interventions, and even simulate potential outcomes of different rehabilitation strategies. Beyond clinical applications, the synthetic EMG data can also find use in the field of human-computer interaction (HCI).

Researchers can leverage the GAN-generated signals to train gesture recognition systems, enabling more intuitive and natu-

ral control of digital interfaces using muscle movements. This can be particularly beneficial for individuals with physical disabilities, providing them with alternative input modalities that enhance their accessibility and independence. Furthermore, the ability to generate synthetic EMG data can have implications in the realm of sports science and performance analysis. Coaches and trainers can use the GAN-generated data to study the biomechanics and muscle activation patterns of elite athletes, helping them develop personalized training programs and optimize athletic performance. Additionally, the synthetic data can be used to simulate various training scenarios, allowing for more efficient and cost-effective evaluations of new training methodologies or equipment.

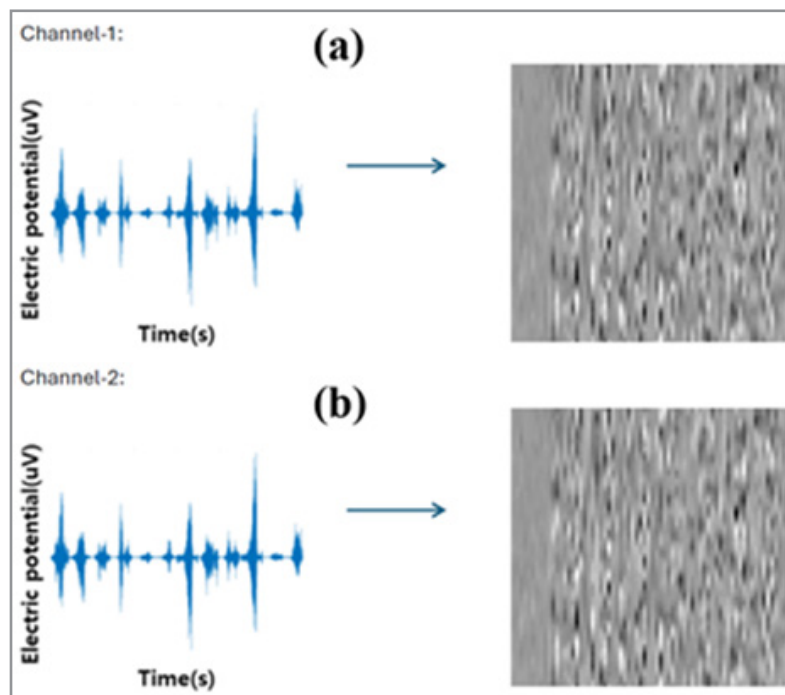


Figure 3: (a) Real 2D Image generated from the original EMG data (b) Artificial 2D Image generated by the trained GAN AI algorithm.

Conclusion

In this study, we demonstrated the feasibility of using Generative Adversarial Networks to generate synthetic EMG data that closely resembles real-world signals. By transforming the original EMG data into a 2D image format and training a GAN model, we were able to produce artificial EMG data that was statistically indistinguishable from the real data. This approach has the potential to significantly streamline the validation and testing of EMG devices, reducing the reliance on extensive data collection from human subjects. Furthermore, the synthetic data can be used to augment existing datasets and improve the performance of EMG-based applications.

Acknowledgment

This research was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE)(2023-RIS009).

References

1. Marinelli, A., Boccardo, N., Semprini, M., Succi, A., Caneppa, M., & et al. (2021). Miniature EMG sensors for prosthetic applications. In 2021 10th International IEEE/EMBS Conference on Neural Engineering (NER) (pp. 1022-1025).
2. de Souza, J. O. de O., Bloedow, M. D., Rubo, F. C., de Figueiredo, R. M., Pessin, G., & et al. (2021). Investigation of different approaches to real-time control of prosthetic hands with electromyography signals. *IEEE Sensors Journal*, 18, 20674-20684.
3. Fleming, A., Stafford, N., Huang, S., Hu, X., Ferris, D. P., & et al. (2021). Myoelectric control of robotic lower limb prostheses: a review of electromyography interfaces, control paradigms, challenges and future directions. *Journal of Neural Engineering*, 4(4), 041004.
4. Sun, Y., Xu, C., Li, G., Xu, W., Kong, J., & et al. (2020). Intelligent human computer interaction based on non-redundant EMG signal. *Alexandria Engineering Journal*, 3, 1149-1157.

5. Mishra, V. K. (2021). A review on emg signal classification for human computer interaction. *Asian Journal of Multidimensional Research*, 12, 681-687.
6. Qureshi, M. F., Mushtaq, Z., ur Rehman, M. Z., Kamavua-ko, E. N. (2022). Spectral image-based multiday surface electromyography classification of hand motions using CNN for human-computer interaction. *IEEE Sensors Journal*, 21, 20676-20683.
7. Taborri, J., Keogh, J., Kos, A., Santuz, A., Umek, A., & et al. (2020). Sport biomechanics applications using inertial, force, and EMG sensors: A literature overview. *Applied Biomechanics and Biomechanics*, 23, 2041549.
8. De Fazio, R., Mastronardi, V. M., De Vittorio, M., Visconti, P., & et al. (2023). Wearable sensors and smart devices to monitor rehabilitation parameters and sports performance: an overview. *Sensors*, 4, 1856.
9. Worsey, M. T. O., Jones, B. S., Cervantes, A., Chauvet, S. P., Thiel, D. V., & et al. (2020). Assessment of head impacts and muscle activity in soccer using a T3 inertial sensor and a portable electromyography (EMG) system: a preliminary study. *Electronics*, 5, 834.
10. Pan, T.-Y., Tsai, W.-L., Chang, C.-Y., Yeh, C.-W., Hu, M.-C., & et al. (2020). A hierarchical hand gesture recognition framework for sports referee training-based EMG and accelerometer sensors. *IEEE Transactions on Cybernetics*, 5, 3172-3183.
11. Ullah, A., Khan, F. S., Mohy-Ud-Din, Z., Hassany, N., Gul, J. Z., & et al. (2024). A Hybrid Approach for Energy Consumption and Improvement in Sensor Network Lifespan in Wireless Sensor Networks. *Sensors*, 5, 1353.
12. Rehman, M. M., Khan, M., ur Rehman, H. M. M., Siddiqui, G. U., Ahmad, Z., & et al. (2024). Nanomaterials in humidity sensors. In *Handbook of Nanomaterials* (Vol. 1, pp. 513-566).
13. Gul, J. Z., Khan, M., Rehman, M. M., Ud Din, Z. M., Kim, W. Y., & et al. (2023). Preparation and Performance Analysis of 3D Thermoformed Fluidic Polymer Temperature Sensors for Aquatic and Terrestrial Applications. *Sensors*, 20, 8506.
14. Khan, M., Rehman, M. M., Khan, S. A., Saqib, M., Kim, W. Y., & et al. (2023). Characterization and performance evaluation of fully biocompatible gelatin-based humidity sensor for health and environmental monitoring. *Frontiers in Materials*, 10, 1233136.
15. Ur Rehman, H. M. M., Khan, M., Rehman, M. M., Khan, S. A., & et al. (2022). High-performance humidity sensor for multipurpose applications by recycling of potato peel bio-waste. *Sensors and Actuators A: Physical*, 343, 113662.
16. Ur Rehman, H. M. M., Rehman, M. M., Saqib, M., Khan, S. A., & et al. (2021). Highly efficient and wide range humidity response of biocompatible egg white thin film. *Nanomaterials*, 7, 1815.
17. Lee, W. (2021). The latest trend in neuromuscular monitoring: return of the electromyography. *Anesthesia and Pain Medicine*, 2, 133-137.
18. Zagrodny, B., Wojnicz, W., Ludwicki, M., Awrejcewicz, J., & et al. (2020). Could thermal imaging supplement surface electromyography measurements for skeletal muscles? *IEEE Transactions on Instrumentation and Measurement*, 70, 1-10.
19. Wong, A. M. H., Furukawa, M., & Maeda, T. (2020). Robustness of rhythmic-based dynamic hand gesture with surface electromyography (sEMG) for authentication. *Electronics*, 12, 2143.
20. Arteaga, M. V., Castiblanco, J. C., Mondragon, I. F., Colorado, J. D., Alvarado-Rojas, C., & et al. (2020). EMG-driven hand model based on the classification of individual finger movements. *Biomedical Signal Processing and Control*, 58, 101834.